



COPY OF PAPER
ORIGINALLY FILED

Attorney's Docket No.: 13425-042001 / 00298-US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Sven Enerbäck et al.
Serial No. : 09/963,285
Filed : September 26, 2001
Title : PROMOTER SEQUENCES

Art Unit : 1645
Examiner :

Commissioner for Patents
Washington, D.C. 20231

Tech Center 1600/2900

#57 MW RECEIVED
JAN 15 2002

TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT UNDER 35 USC §119

In accordance with the provisions of 35 U.S.C. §119, applicants hereby claim priority of Swedish Patent application No. 0003435-5, filed September 26, 2000. A certified copy of the application is submitted herewith. As the priority application is in the English language, all of the requirements of §119 have been met.

Please apply any charges or credits to Deposit Account No. 06-1050, referencing Attorney Docket No. 13425-042001.

Respectfully submitted,

Date: November 9, 2001

Jack Brennan
Jack Brennan
Reg. No. 47,443

Fish & Richardson P.C.
225 Franklin Street
Boston, Massachusetts 02110-2804
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

20331720.doc

RECEIVED
MARCH 26 2001
TECH CENTER 1600
U.S. PATENT AND TRADEMARK OFFICE

CERTIFICATE OF MAILING BY FIRST CLASS MAIL

I hereby certify under 37 CFR §1.8(a) that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

November 9, 2001
Date of Deposit

Darlene J. Morin
Signature

Darlene J. Morin
Typed or Printed Name of Person Signing Certificate

PRV

PATENT- OCH REGISTRERINGSVERKET
Patentavdelningen



COPY OF PAPER
ORIGINALLY FILED

Intyg
Certifikat

Härmed intygas att bifogade kopior överensstämmer med de handlingar som ursprungligen ingivits till Patent- och registreringsverket i nedannämnda ansökan.

This is to certify that the annexed is a true copy of the documents as originally filed with the Patent- and Registration Office in connection with the following patent application.

(71) Sökande Pharmacia & Upjohn AB, Stockholm SE
Applicant (s)

(21) Patentansökningsnummer 0003435-5
Patent application number

(86) Ingivningsdatum 2000-09-26
Date of filing

Stockholm, 2001-09-19

För Patent- och registreringsverket
For the Patent- and Registration Office

Kerstin Gerdén
Kerstin Gerdén

Avgift
Fee 170:-

RECEIVED
JAN 15 2002
TECH CENTER 1600/2900 MAR 26 2002
RECEIVED
TECH CENTER 1600/2900



PRU0000926

RECEIVED
JAN 15 2002
TECH CENTER 1600/2300RECEIVED
JAN 15 2002
TECH CENTER 1600/2300

PROMOTER SEQUENCES

TECHNICAL FIELD

The present invention relates an isolated promoter region of the mammalian transcription factor *FOXC2*. The invention also relates to screening methods for agents modulating the expression of *FOXC2* and thereby being potentially useful for the treatment of medical conditions related to obesity. The invention further relates to a previously unknown variant of the human *FOXC2* gene, derived via the use of an alternative promoter, which produces an additional exon that generates a distinct open reading frame via splicing. The alternative gene encodes a variant of the *FOXC2* transcription factor, which is lacking a part of the DNA-binding domain and consequently has a potential regulatory function.

BACKGROUND ART

More than half of the men and women in the United States, 30 years of age and older, are now considered overweight, and nearly one-quarter are clinically obese. This high prevalence has led to increases in the medical conditions that often accompany obesity, especially non-insulin dependent diabetes mellitus (NIDDM), hypertension, cardiovascular disorders, and certain cancers. Obesity results from a chronic imbalance between energy intake (feeding) and energy expenditure. To better understand the mechanisms that lead to obesity and to develop strategies in certain patient populations to control obesity, there is a need to develop a better underlying knowledge of the molecular events that regulate the differentiation of preadipocytes and stem cells to adipocytes, the major component of adipose tissue.

The helix-loop-helix (HLH) family of transcriptional regulatory proteins are key players in a wide array of developmental processes (for a review, see Massari & Murre (2000) Mol. Cell. Biol. 20: 429-440). Over 240 HLH proteins have been identified to date in organisms ranging from the yeast *Saccharomyces cerevisiae* to humans. Studies in *Xenopus laevis*, *Drosophila melanogaster*, and mice have convincingly demonstrated that HLH proteins are intimately involved in developmental events such as cellular differentiation, lineage

RECEIVED
MAR 26 2002
TECH CENTER 1600/2300

commitment, and sex determination. In multicellular organisms, HLH factors are required for a multitude of important developmental processes, including neurogenesis, myogenesis, hematopoiesis, and pancreatic development.

- 5 The winged helix / forkhead class of transcription factors is characterized by a 100-amino acid, monomeric DNA-binding domain. X-ray crystallography of the forkhead domain from HNF-3 γ has revealed a three-dimensional structure, the "winged helix", in which two loops (wings) are connected on the C-terminal side of the helix-loop-helix (for reviews, see
Brennan, R.G. (1993) Cell 74: 773-776; and Lai, E. et al. (1993) Proc. Natl. Acad. Sci.
10 U.S.A. 90: 10421-10423).

The isolation of the mouse mesenchyme forkhead-1 (MFH-1) and the corresponding human (*FKHL14*) chromosomal genes is disclosed by Miura, N. et al. (1993) FEBS letters 326: 171-176; and (1997) Genomics 41: 489-492. The nucleotide sequences of the mouse
15 MFH-1 gene and the human *FKHL14* gene have been deposited with the EMBL/GenBank Data Libraries under accession Nos. Y08222 (SEQ ID NO: 5) and Y08223 (SEQ ID NO: 8), respectively. A corresponding gene has been identified in *Gallus gallus* (GenBank accession numbers U37273 and U95823).

- 20 The International Patent Application WO 98/54216 discloses a gene encoding a Forkhead-Related Activator (FREAC)-11 (also known as S12), which is identical with the polypeptide encoded by the human *FKHL14* gene disclosed by Miura, *supra*. This transcription factor is expressed in adipose tissue and involved in lipid metabolism and adipocyte differentiation (cf. Swedish patent application No. 0000531-4, filed February 18,
25 2000).

The nomenclature for the winged helix / forkhead transcription factors has been standardized and Fox (Forkhead Box) has been adopted as the unified symbol (Kaestner et al. (2000) Genes & Development 14: 142-146; see also <http://www.biology.pomona.edu/fox>). It has been agreed that the genes previously designated MFH-1 and *FKHL14* (as well as FREAC-11 and S12) should be designated *FOXC2*.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the general structure of the human *FOXC2* gene.

- 5 Figure 2 illustrates the results from phylogenetic footprinting experiments. Shown is the fraction conserved (1.0 = 100%) between mouse FoxC2 and human *FOXC2* sequences in the alignment generated with Clustal. Solid (bold) line indicates the fraction of the human sequence which is identical to the mouse within a 200 bp "window" over the human sequence in the alignment. The weak (dotted) line is set to -0.05 when the sliding window
10 contains human exon sequence and to -0.1 when the window is entirely composed of exon sequence. Regions containing local maxima or exceeding a conservation fraction of 0.7 are likely to be functional and are classified as "predicted regulatory regions".

- 15 Figure 3 illustrates the predicted "enhancer" region in the human *FOXC2* gene. Underlined sequences indicate likely transcription factor binding sites. Boxed sequence indicates exon sequence.

Splice = sequence predicted as splice site in the alternatively spliced gene;
E-box-like = sequence resembling the "E-box" motif CANNTG known as a target for DNA binding proteins containing a helix-loop-helix domain (often associated with the activation
20 of cell-type specific gene transcription during tissue differentiation; see Massari & Murre (2000) Mol. Cell. Biol. 20: 429-440)

Forkhead-like = sequence resembling binding site for the winged helix / forkhead class of transcription factors;

Ets-like = sequence resembling consensus binding site for ETS-domain transcription factor family (see Sharrocks et al. (1997) Int. J. Biochem. Cell Biol. 29, 1371-1387).

Figure 4 illustrates the predicted "promoter" region in the human *FOXC2* gene. Underlined sequence indicates exon sequences. Boxed sequences indicate conserved block (potential transcription factor binding sites).

DESCRIPTION OF THE INVENTION

According to the present invention, the partially known sequence (SEQ ID NO: 8) of human *FOXC2* gene has been extended. In the previously unknown region of the gene, 5 differentially conserved regions, consistent with regulatory function, have been identified. Further, an alternative transcript has been identified, which includes the use of at least two exons. The putative regulatory enhancer is immediately adjacent to the newly discovered alternative exon, suggesting that it may play a role in the alternative selection of transcript classes.

10

Modulation of the *FOXC2* regulation is expected to have therapeutic value in type II diabetes; obesity, hypercholesterolemia, and other cardiovascular diseases or dyslipidemias.

15 Consequently, in a first aspect this invention provides a human *FOXC2* promoter region comprising a sequence selected from:

(a) the nucleotide sequence set forth as positions 1250 to 2235, such as positions 1250 to 1749 or positions 1692 to 1703, in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* promoter activity;

20 (b) the complementary strand of (a); and

(c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).

Another aspect of the invention is a recombinant construct comprising the human *FOXC2* 25 promoter region as defined above. In the said recombinant construct, the human *FOXC2* promoter region can be operably linked to a gene encoding a detectable product, such as the human *FOXC2* gene, or a reporter gene. The term "operably linked" as used herein means functionally fusing a promoter with a structural gene in the proper frame to express the structural gene under control of the promoter. As used herein, the term "reporter gene"

30 means a gene encoding a gene product that can be identified using simple, inexpensive methods or reagents and that can be operably linked to the human *FOXC2* promoter region or an active fragment thereof. Reporter genes such as, for example, a luciferase, β -galactosidase, alkaline phosphatase, or green fluorescent protein reporter gene, can be used

to determine transcriptional activity in screening assays according to the invention (see, for example, Goeddel (ed.), *Methods Enzymol.*, Vol. 185, San Diego: Academic Press, Inc. (1990); see also Sambrook, *supra*).

- 5 The invention also provides a vector comprising the recombinant construct as defined above, as well as a host cell stably transformed with such a vector, or generally with the recombinant construct according to the invention. The term "vector" refers to any carrier of exogenous DNA that is useful for transferring the DNA to a host cell for replication and/or appropriate expression of the exogenous DNA by the host cell.

10

- In another aspect, the invention provides a method for identification of an agent regulating *FOXC2* promoter activity, said method comprising the steps: (i) contacting a candidate agent with a human *FOXC2* promoter region as defined above; and (ii) determining whether said candidate agent modulates expression of the *FOXC2* gene, such modulation being indicative for an agent capable of regulating *FOXC2* promoter activity. As used herein, the term "agent" means a biological or chemical compound such as a simple or complex organic molecule, a peptide, a protein or an oligonucleotide.

- A transfection assay can be a particularly useful screening assay for identifying an effective agent. In a transfection assay, a nucleic acid containing a gene such as a reporter gene that is operably linked to a human *FOXC2* promoter, or an active fragment thereof, is transfected into the desired cell type. A test level of reporter gene expression is assayed in the presence of a candidate agent and compared to a control level of expression. An effective agent is identified as an agent that results in a test level of expression that is different than a control level of reporter gene expression, which is the level of expression determined in the absence of the agent. Methods for transfecting cells and a variety of convenient reporter genes are well known in the art (see, for example, Goeddel (ed.), *Methods Enzymol.*, Vol. 185, San Diego: Academic Press, Inc. (1990); see also Sambrook, *supra*). Consequently, the said method could e.g. comprising assaying reporter gene expression in a host cell, stably transformed with a recombinant construct comprising the human *FOXC2* promoter, in the presence and absence of a candidate agent, wherein an effect on the test level of expression as compared to control level of expression is indicative of an agent capable of regulating *FOXC2* promoter activity.

In a further aspect, the invention provides a human *FOXC2* enhancer region comprising a sequence selected from:

- (a) the nucleotide sequence set forth as positions 216 to 475, such as positions 223 to 5 231, positions 359 to 375, positions 378 to 402, or positions 403 to 423, in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
- (b) the complementary strand of (a); and
- (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).

10

As described above for the human *FOXC2* promoter region, the invention further provides a recombinant construct comprising a human *FOXC2* enhancer region, a vector comprising the said recombinant construct, as well as a host cell stably transformed with said vector or with said recombinant construct.

15

Further, the invention provides a method for identification of an agent regulating *FOXC2* enhancer activity, said method comprising the steps: (i) contacting a candidate agent with the human *FOXC2* enhancer region as defined above; and (ii) determining whether said candidate agent modulates expression of the *FOXC2* gene, such modulation being indicative for an agent capable of regulating *FOXC2* enhancer activity. It will be understood by the skilled person that known steps are available for performing such a method. For instance, a "panel" of constructs which include a variety of mutations and deletions can be used in order to associate a response with a specific alteration of a single base or subsegment of the regulatory apparatus. A simple panel might include: enhancer 20 plus promoter, promoter only, enhancer plus a "minimal" promoter from a distinct gene. As mentioned above, a transfection assay, using a host cell stably transformed with a suitable recombinant construct, can be a particularly useful screening assay for identifying 25 an effective agent.

30 In yet a further aspect, the invention provides a method for identification of an agent capable of regulating a mammalian *FOXC2* promoter activity, said method comprising the steps (i) contacting a candidate agent with a murine *FoxC2* promoter nucleotide sequence shown as positions 216 to 2235, such as positions 216 to 475 or positions 1250 to 2235, in

SEQ ID NO: 5; and (ii) determining whether said candidate agent modulates expression of a mammalian *FOXC2* gene, such modulation being indicative for an agent capable of regulating mammalian *FOXC2* promoter activity.

- 5 In another important aspect, the invention provides an isolated nucleic acid molecule selected from:
- (a) nucleic acid molecules comprising a nucleotide sequence as shown in SEQ ID NO: 3;
 - (b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the
 - 10 polypeptide coding region of a nucleic acid molecule as defined in (a) and which codes for a variant form of the *FOXC2* transcription factor; and
 - (c) nucleic acid molecules comprising a nucleic acid sequence which is degenerate as a result of the genetic code to a nucleotide sequence as defined in (a) or (b) and which codes for a variant form of the *FOXC2* transcription factor.

15 In a preferred form of the invention, the said nucleic acid molecule has a nucleotide sequence identical with SEQ ID NO: 3 of the Sequence Listing. However, the nucleic acid molecule according to the invention is not to be limited strictly to the sequence shown as SEQ ID NO: 3. Rather the invention encompasses nucleic acid molecules carrying modifications like substitutions, small deletions, insertions or inversions, which nevertheless encode proteins having substantially the biochemical activity of the *FOXC2* polypeptide according to the invention. Included in the invention are consequently nucleic acid molecules, the nucleotide sequence of which is at least 90% homologous, preferably at least 95% homologous, with the nucleotide sequence shown as SEQ ID NO: 3 in the

20 Sequence Listing.

Included in the invention is also a nucleic acid molecule which nucleotide sequence is degenerate, because of the genetic code, to the nucleotide sequence shown as SEQ ID NO: 3. A sequential grouping of three nucleotides, a "codon", codes for one amino acid. Since

30 there are 64 possible codons, but only 20 natural amino acids, most amino acids are coded for by more than one codon. This natural "degeneracy", or "redundancy", of the genetic code is well known in the art. It will thus be appreciated that the nucleotide sequence

shown in the Sequence Listing is only an example within a large but definite group of sequences which will encode the variant FOXC2 polypeptide.

The invention includes an isolated polypeptide encoded by the nucleic acid as defined above. In a preferred form, the said polypeptide has an amino acid sequence according to SEQ ID NO: 4 of the Sequence Listing. However, the polypeptide according to the invention is not to be limited strictly to a polypeptide with an amino acid sequence identical with SEQ ID NO: 4 in the Sequence Listing. Rather the invention encompasses polypeptides carrying modifications like substitutions, small deletions, insertions or inversions, which polypeptides nevertheless have substantially the biological activities of the variant FOXC2 polypeptide.

A further aspect of the invention is a vector harboring the nucleic acid molecule according to the invention. The said vector can e.g. be a replicable expression vector, which carries and is capable of mediating the expression of a DNA molecule according to the invention.

In the present context the term "replicable" means that the vector is able to replicate in a given type of host cell into which it has been introduced. Examples of vectors are viruses such as bacteriophages, cosmids, plasmids and other recombination vectors. Nucleic acid molecules are inserted into vector genomes by methods well known in the art.

Included in the invention is also a cultured host cell harboring a vector according to the invention. Such a host cell can be a prokaryotic cell, a unicellular eukaryotic cell or a cell derived from a multicellular organism. The host cell can thus e.g. be a bacterial cell such as an *E. coli* cell; a cell from yeast such as *Saccharomyces cerevisiae* or *Pichia pastoris*, or a mammalian cell. The methods employed to effect introduction of the vector into the host cell are standard methods well known to a person familiar with recombinant DNA methods.

In yet another aspect, the invention includes a method for identifying an agent capable of regulating expression of the nucleic acid molecule as defined above, said method comprising the steps (i) contacting a candidate agent with the said nucleic acid molecule; and (ii) determining whether said candidate agent modulates expression of the said nucleic acid molecule.

In another aspect the invention provides an antisense oligonucleotide having a sequence capable of specifically hybridizing to RNA transcribed by the alternatively spliced nucleic acid molecule shown as SEQ ID NO: 3, so as to prevent translation of the said RNA.

- 5 Antisense nucleic acids (preferably 10 to 20 base-pair oligonucleotides) capable of specifically binding to control sequences for the alternatively spliced *FOXC2* gene are introduced into cells, e.g. by a viral vector or colloidal dispersion system such as a liposome. The antisense nucleic acid binds to the target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and
- 10 methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. Suppression of expression of the alternatively spliced *FOXC2* gene, at either the transcriptional or translational level, is useful to generate cellular or animal models for diseases/conditions related to lipid metabolism.
- 15 Throughout this description the terms "standard protocols" and "standard procedures", when used in the context of molecular biology techniques, are to be understood as protocols and procedures found in an ordinary laboratory manual such as: Current Protocols in Molecular Biology, editors F. Ausubel et al., John Wiley and Sons, Inc. 1994, or Sambrook, J., Fritsch, E.F. and Maniatis, T., Molecular Cloning: A laboratory manual,
- 20 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 1989.

EXAMPLES

- 25 Additional features of the invention will be apparent from the following Examples. Examples 1 to 4 are actual, while the remaining Examples are prophetic.

EXAMPLE 1: Computational identification of *FOXC2* genomic sequences

- 30 The sequences present in the GenBank database (<http://www.ncbi.nlm.nih.gov>) were screened for sequence similarity to the human *FOXC2* cDNA sequence (GenBank accession number NM_00521 (SEQ ID NO: 9)). The BLAST algorithm (Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402) was used for determining sequence identity.

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). A working draft genomic sequence in 25 unordered pieces, from the *Homo sapiens* chromosome 16 clone RP11-463O9 (GenBank accession number AC009108; Version 6; GI:7689930; released 4 May 5 2000), was selected for further studies.

Regions in sequence AC009108 matching portions of the *FOXC2* cDNA sequence NM_005251 were combined using the PHRAP software, developed at the University of Washington (<http://www.genome.washington.edu/UWGC/analysistools/phrap.htm>). Two 10 contigs of 9780 bp (positions 116445 to 126224 in GenBank AC009108.6) and 3784 bp (positions 42927 to 46710 in GenBank AC0091108.6), respectively, were assembled to generate a human *FOXC2* genomic fragment of 13451 bp.

The ClustalW multiple sequence alignment program, version 1.8 (Thompson et al. (1994) 15 Nucleic Acids Research 22: 4673-4680), was then used to identify the human *FOXC2* extended genomic DNA sequence of 6458 bp (SEQ ID NO: 1) by comparison with the mouse cDNA sequence X74040 (SEQ ID NO: 6). First, a 6459 bp sequence, corresponding to positions 1500-7958 in the 13451 bp sequence, was selected. Positions 1-2285 in this 6459 bp sequence corresponded to 44426-46710 in AC009108.6, while positions 2151-20 6459 corresponded to positions 126224-121916 (reverse complement taken) in AC009108.6. The overlap of positions 2151-2285 allowed for the contigs to be joined by the assembly program. The G residue in position 2655 was considered to be a sequencing error and was removed, which resulted in the 6458 bp sequence set forth as SEQ ID NO: 1. The open reading frame in SEQ ID NO: 1 encodes a polypeptide (SEQ ID NO: 2) identical 25 with the known human *FOXC2* polypeptide shown as SEQ ID NO:10.

EXAMPLE 2: Identification of potential regulatory sequences in the human and mouse *FOXC2* genomic sequences

In phylogenetic footprinting (for a review, see Duret & Bucher (1997) Current Opinion in Structural Biology 7(3): 399-406) sequences are aligned and a regional sequence identity is determined for each window of a fixed, arbitrary length. This allows the identification of 30

- 11 -

potential regulatory regions in genomic sequences. Non-exon sequences that are conserved over the course of evolution are likely to perform regulatory roles. Phylogenetic footprinting was performed as described in Wasserman & Fickett (1998) J. Mol. Biol. 278, 167-181, based on an alignment generated with the ClustalW multiple sequence alignment program, version 1.8 (Thompson et al. (1994) Nucleic Acids Research 22: 4673-4680), with default parameters adjusted to a gap opening penalty of 20 and a gap extension penalty of 0.2. The human (SEQ ID NO: 1) and mouse (SEQ ID NO: 5) genomic sequences were aligned. Percentage identity was plotted for each contiguous 200 bp segment of the human gene to identify segments differentially conserved (in comparison to adjoining sequences) (Fig. 2).

In addition to segments of the published exon sequence, two differentially conserved regions or "footprints" were identified in the human gene. Both of these regions are local maxima and contain segments which exceed 70% nucleotide identity between the human and mouse genomic sequences. One region, shown as positions 1250 to 2235, in particular positions 1250 to 1749, in SEQ ID NO: 1, immediately adjacent to the published exon region, is likely to contain the transcription start site and proximal promoter regulatory sequences (Fig. 4). Another region, shown as positions 216 to 475 in SEQ ID NO: 1, approximately 1700 bp distal from the transcription start site, is likely to function as some form of regulatory region (either enhancer or repressor) (Fig. 3). (A schematic overview of the extended *FOXC2* gene is shown in Fig. 1.)

Further analysis of these regulatory regions identified short segments of higher conservation between the mouse and human genes, suggesting that these specific segments function as transcription factor binding sites. Screening of the TRANSFAC transcription factor database (<http://transfac.gbf.de>) (see Wingender et al. (2000) Nucleic Acids Research 28(1): 316-319) for matches to known transcription factors suggested the presence of multiple forkhead-like binding sites in the distal regulatory enhancer, which suggests potential auto-regulation of *FOXC2* by its protein product.

A third region containing a local maxima for conservation and a segment exceeding 70% identity is present at the 3' end of the published exon sequence within the 3' UTR. This conserved region may have a role in mRNA processing.

The same analysis was performed with reference to 200 bp contiguous segments of the mouse *FoxC2* genomic sequence (SEQ ID NO: 5). The following conserved regions were identified: 190 to 420; 1070 to 1645; and 5580 to 5875. They correlate to the regions
5 indicated above for the human sequence and should be considered orthologous regions.

EXAMPLE 3: Identification of an alternative human *FOXC2* cDNA sequence

10 BLASTN screening of the dbEST database from GenBank, using the human *FOXC2* cDNA (SEQ ID NO: 9) as a query sequence, revealed several ESTs overlapping containing portions of the available cDNA. A specialized tool, est_genome (<http://www.sanger.ac.uk>), for the prediction of exon boundaries using ESTs was applied to compare the EST sequences to the genomic sequences (See Mott, R. (1997) Computer Applications in the
15 Biosciences 13(4): 477-478). Two classes of ESTs were observed: sequences extending into the 3'-untranslated region and sequences revealing an alternative first exon spliced to a junction internal to the previously described first exon.

Specifically, it was found that the nucleotides in positions 33 to 182 in the EST with
20 accession no. AW271272 (SEQ ID NO: 11) were identical to positions 66 to 215 in the extended *FOXC2* genomic sequence (SEQ ID NO: 1), and that positions 183 to 327 in SEQ ID NO: 11 were identical to positions 2516 to 2660 in SEQ ID NO: 1. Similarly, positions 5 to 55 in the EST with accession no. AW793237 (SEQ ID NO: 12) were identical to positions 165 to 215 in the extended *FOXC2* genomic sequence (SEQ ID NO:
25 1), and positions 56 to 157 in SEQ ID NO: 12 were identical to positions 2516 to 2607 in SEQ ID NO: 1. These results revealed an alternative splicing pattern in the human *FOXC2* gene. According to this splicing pattern, an alternative gene sequence (SEQ ID NO: 3) is derived by joining the regions shown as positions 1-215 and 2516-6458 in SEQ ID NO: 1. Alternative splicing patterns are known to regulate the synthesis of a variety of peptides
30 and proteins. It may result in proteins with an entirely different function or in dysfunctional or inhibitory splice products (for a review, see McKeown (1992) Annu. Rev. Cell. Biol. 8: 133-155).

The amino acids corresponding to positions 1 to 94 in the published FOXC2 transcription factor (SEQ ID NO: 10) are missing in protein encoded by the spliced variant generated from the alternative promoter (SEQ ID NO: 4). Consequently, the entire region N-terminal of the DNA binding domain and a portion of the DNA-binding domain (corresponding to 5 positions 72-94 in SEQ ID NO: 2) are not present in the splice variant. It is postulated that this truncation leads to a protein which has a deficient "forkhead" DNA-binding region, and thus has a potential inhibitory function on the biological activities of the FOXC2 protein. This truncated FOXC2 protein may have a role in regulation of FOXC2, and an involvement in adipocyte differentiation and adipogenesis.

10

EXAMPLE 4: Cloning and sequencing of the FOXC2 promoter

The DNA region corresponding to nucleotide 176 to nucleotide 2233 (SEQ ID NO. 1 15 version 2) has been cloned using nested PCR on human genomic DNA. The PCR was performed according the Herculase™ protocol (Stratagene catalog #600260; <http://www.stratagene.com/pcr/herculase.htm>) and with the inclusion of 8-10% DMSO.

In the initial reaction, the 5'-primer KRKX131 (CCATTGCCTTCTAGTCGCCTCC) was 20 used together with the 3'-primer KRKX133 (CGTTGGGGTCGGACACGGAGTA) using 250 ng Clontech Genomic DNA # 6550-1 as template. The nested reaction was performed on 1/100 of the initial PCR reaction using the 5'-primer KRKX132 (GGTACCTACGCAGCCGATGAACAGCCA) and the 3'-primer KRKX134 (GCTAGCGCTGCTTCCGAGACGGCTCG). After the second PCR, the product was 25 analyzed by electrophoresis in a 1.2% agarose gel, and a PCR product of the expected size was obtained and extracted for ligation into a TOPO PCR2.1 vector (Invitrogen, Carlsbad, CA) by standard cloning procedures and thereafter sequenced. The PCR reaction and cloning procedure was repeated in two parallel separate experiments, and sequence data from the two separate reactions were compared with the bioinformatically assembled 30 sequence.

A DNA region containing the promoter (Fig. 4) corresponding to nt1179 to 2233 (SEQ ID NO: 1, version 2) was has been cloned using nested PCR in the same manner as described

above. In the initial reaction, the 5'-primer KRKX136 (GGTACCCCCCGAGCCTGGAAACTCCCT) was used together with the 3'-primer KRKX134 (GCTAGCGCTGCTTCCGAGACGGCTCG) using 250 ng genomic DNA as a template. The PCR reaction and cloning procedure was repeated in four parallel separate 5 experiments, and sequence data from the four separate reactions were compared with the bioinformatically assembled sequence.

EXAMPLE 5: Tissue expression profiling of the alternative transcript

10

Tissue expression profiling of the alternative transcript (SEQ ID NO: 3) is performed using standard Northern blotting procedures. RNA samples from an array of human tissues, including adipose tissue, are analyzed using an RNA or DNA probe specific for the alternative transcript. The expression profile in adipose tissue could be indicative a 15 putative regulatory function for the alternative gene product (SEQ ID NO: 4) on adipogenesis and adipocyte differentiation.

In addition, reverse transcriptase PCR (RT-PCR) according to standard procedures is used to detect very low level expression of the alternative transcript in adipose tissue. RNA is 20 prepared from human adipose tissue, and RT-PCR is performed using PCR primers specific for the alternative transcript.

EXAMPLE 6: Mapping of the 5'-edge of the alternative exon by RACE-PCR

25

RNA is prepared from human adipose tissue using standard protocols. RACE (Rapid Amplification of cDNA Ends) PCR is performed using the SMART™ RACE cDNA Amplification Kit (Clontech catalogue No. K1811-1; <http://www.clontech.com/product/catalog/PCR/smartrace.html>). With this procedure, the first strand synthesis produces 30 cDNA with an extension containing a known sequence. Due to the mechanism of the extension procedure, the extension is typically added only to complete first strand cDNAs. The 5'-RACE PCR is then performed using the 5'-primer provided with the kit, together with a 3'-primer corresponding to positions 210-237 in SEQ ID NO: 3

(GAACTGGTAGATGCCGTTCAAGGTTCC) specific for the alternative transcript. The PCR product is cloned into a cloning vector and sequenced using standard protocols.

5 EXAMPLE 7: Functional analysis

The identified regulatory regions are analyzed to determine their impact on the transcription of the *FOXC2* gene or a reporter gene substituted for *FOXC2*. A PCR reaction is performed to isolate the promoter region adjacent to the published exon sequence, possibly including the sequences extending to the beginning of the ATG encoding the first methionine. This PCR product is cloned into a reporter plasmid adjacent to a reporter gene (e.g. luciferase). The upstream regulatory region, i.e. regions containing both upstream and promoter proximal sequences, or these sequences bearing artificially induced differences, are cloned in a similar manner. These constructs are transfected into a cell culture model system and the level/activity of the protein encoded by the reporter gene is determined. This would provide information on the function of the identified regions, and used to assess the impact of the different regions on transcriptional regulation. Similarly, the upstream regulatory region, a region containing both upstream and promoter proximal sequences, or these sequences bearing artificially induced differences can be cloned and used to assess the impact of these regions on the transcription of the reporter gene.

25 EXAMPLE 8: Reporter gene assay to identify modulating compounds

Reporter gene assays are well known as tools to signal transcriptional activity in cells. (For a review of chemiluminescent and bioluminescent reporter gene assays, see Bronstein et al. (1994) Analytical Biochemistry 219, 169-181.) For instance, the photoprotein luciferase provides a useful tool for assaying for modulators of promoter activity. Cells are transiently transfected with a reporter construct which includes a gene for the luciferase protein downstream from the *FOXC2* promoter and enhancer region, or fragments thereof regulating the *FOXC2* activity. Luciferase activity may be quantitatively measured using e.g. luciferase assay reagents that are commercially available from Promega (Madison,

- 16 -

WI). Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity.

60
100
200
300
400
500

TABLE I
Summary of *FOXC2* sequences

SEQ ID NO:	GenBank accession no.	Description
1		Human <i>FOXC2</i> extended genomic DNA sequence
2		Human <i>FOXC2</i> polypeptide sequence (Identical with SEQ ID NO: 10)
3		Human <i>FOXC2</i> DNA sequence Alternative splicing
4		Human polypeptide sequence Alternative open reading frame
5	Y08222	Mouse MHF-1 (<i>FoxC2</i>) genomic DNA sequence (CDS 2070 – 3554)
6	X74040	Mouse MHF-1 (<i>FoxC2</i>) cDNA sequence
7		Mouse MHF-1 (<i>FoxC2</i>) polypeptide sequence
8	Y08223	Human <i>FKHL14</i> (<i>FOXC2</i>) genomic DNA sequence (CDS 1197 – 2702)
9	NM_005251	Human <i>FKHL14</i> (<i>FOXC2</i>) cDNA sequence
10		Human <i>FKHL14</i> (<i>FOXC2</i>) polypeptide sequence
11	AW 271272	Human EST
12	AW 793237	Human EST

TABLE II

Summary of features in human *FOXC2* sequences shown as SEQ ID NOs: 1 and 3

Feature	Positions
SEQ ID NO: 1	
First exon according to the alternative transcript	1 – 215
– Untranslated region	1 – 186
– Region coding for 5'-part of alternative protein	187 – 215
Alternative first exon splice site	215 – 216
Predicted enhancer region	216 – 475
– E-box-like region	223 – 231
– Forkhead-like region	359 – 375
– Forkhead-like region	378 – 402
– Ets-like region	403 – 423
Predicted promoter region	1250 – 1749
– Forkhead-like region	1692 – 1703
First exon according to the published form of the transcript	1746 – 4629
– Untranslated region	1746 – 2234
– Polypeptide coding region	2235 – 3740
– Region coding for DNA-binding domain	2448 – 2735
Second exon according to the alternative transcript	2516 – 4629
– Portion of polypeptide used in alternative transcript	2516 – 3740
– Untranslated region	3741 – 4629
SEQ ID NO: 3	
Polypeptide coding region (5' of splice site)	187 – 215
Polypeptide coding region (3' of splice site)	216 – 1437
– Region coding for truncated portion of protein	216 – 435

CLAIMS

1. A human *FOXC2* promoter region comprising a sequence selected from:
 - (a) the nucleotide sequence set forth as positions 1692 to 1703 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* promoter activity;
 - (b) the complementary strand of (a); and
 - (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
- 10 2. The human *FOXC2* promoter region according to claim 1, comprising a sequence selected from:
 - (a) the nucleotide sequence set forth as positions 1250 to 1749 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* promoter activity;
 - (b) the complementary strand of (a); and
 - (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
- 15 3. The human *FOXC2* promoter region according to claim 2, comprising a sequence selected from:
 - (a) the nucleotide sequence set forth as positions 1250 to 2235 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* promoter activity;
 - (b) the complementary strand of (a); and
 - (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
- 20 4. A recombinant construct comprising the human *FOXC2* promoter region according to any one of claims 1 to 3.
- 25 5. The recombinant construct according to claim 4 wherein the human *FOXC2* promoter region is operably linked to a gene encoding a detectable product.
- 30 6. The recombinant construct according to claim 5 wherein said gene encoding a detectable product is a human *FOXC2* gene.

7. The recombinant construct according to claim 4 further comprising a reporter gene.
8. A vector comprising the recombinant construct according to any one of claims 4 to 7.
9. A host cell stably transformed with the recombinant construct according to any one of claims 4 to 7.
10. A method for identification of an agent regulating *FOXC2* promoter activity, said method comprising the steps
 - (i) contacting a candidate agent with a human *FOXC2* promoter region as defined in any one of claims 1 to 3; and
 - (ii) determining whether said candidate agent modulates expression of the *FOXC2* gene, such modulation being indicative for an agent capable of regulating *FOXC2* promoter activity.
11. A method for identification of an agent regulating *FOXC2* promoter activity, said method comprising assaying reporter gene expression in a cell according to claim 9 in the presence and absence of a candidate agent, wherein an effect on the test level of expression as compared to control level of expression is indicative of an agent capable of regulating *FOXC2* promoter activity.
12. A human *FOXC2* enhancer region comprising a sequence selected from:
 - (a) the nucleotide sequence set forth as positions 223 to 231 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
 - (b) the complementary strand of (a); and
 - (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
13. A human *FOXC2* enhancer region comprising a sequence selected from:
 - (a) the nucleotide sequence set forth as positions 359 to 375 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
 - (b) the complementary strand of (a); and

- (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
14. A human *FOXC2* enhancer region comprising a sequence selected from:
- 5 (a) the nucleotide sequence set forth as positions 378 to 402 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
- (b) the complementary strand of (a); and
- (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
- 10 15. A human *FOXC2* enhancer region comprising a sequence selected from:
- (a) the nucleotide sequence set forth as positions 403 to 423 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
- (b) the complementary strand of (a); and
- (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
16. The human *FOXC2* enhancer region according to any one of claims 12 to 15 comprising a sequence selected from:
- 20 (a) the nucleotide sequence set forth as positions 216 to 475 in SEQ ID NO: 1, or a fragment thereof exhibiting *FOXC2* enhancer activity;
- (b) the complementary strand of (a); and
- (c) nucleotide sequences capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence as defined in (a) or (b).
- 25 17. A recombinant construct comprising a human *FOXC2* enhancer region according to any one of claims 12 to 15.
18. A vector comprising the recombinant construct according to claim 17.
- 30 19. A host cell stably transformed with the recombinant construct according to claim 18.

20. A method for identification of an agent regulating *FOXC2* enhancer activity, said method comprising the steps
 - (i) contacting a candidate agent with the human *FOXC2* enhancer region as defined in any one of claims 12 to 16; and
 - 5 (ii) determining whether said candidate agent modulates expression of the *FOXC2* gene, such modulation being indicative for an agent capable of regulating *FOXC2* enhancer activity.
- 10 21. A method for identification of an agent capable of regulating *FOXC2* enhancer activity, said method comprising assaying reporter gene expression in a cell as defined in claim 19 in the presence and absence of a candidate agent, wherein an effect on the test level of expression as compared to control level of expression is indicative of an agent capable of regulating *FOXC2* enhancer activity.
- 15 22. A method for identification of an agent capable of regulating a mammalian *FOXC2* promoter activity, said method comprising the steps
 - (i) contacting a candidate agent with a murine *FoxC2* promoter nucleotide sequence shown as positions 1250 to 2235 in SEQ ID NO: 5; and
 - (ii) determining whether said candidate agent modulates expression of a mammalian *FOXC2* gene, such modulation being indicative for an agent capable of regulating mammalian *FOXC2* promoter activity.
- 20 23. A method for identification of an agent capable of regulating a mammalian *FOXC2* enhancer activity, said method comprising the steps
 - (i) contacting a candidate agent with a murine *FoxC2* enhancer nucleotide sequence shown as positions 216 to 475 in SEQ ID NO: 5; and
 - (ii) determining whether said candidate agent modulates expression of a mammalian *FOXC2* gene, such modulation being indicative for an agent capable of regulating mammalian *FOXC2* enhancer activity.
- 25 30 24. A method for identification of an agent capable of regulating a mammalian *FOXC2* enhancer activity, said method comprising the steps

- (i) contacting a candidate agent with a murine *FoxC2* enhancer nucleotide sequence shown as positions 216 to 2235 in SEQ ID NO: 5; and
(ii) determining whether said candidate agent modulates expression of a mammalian *FOXC2* gene, such modulation being indicative for an agent capable of regulating
5 mammalian *FOXC2* enhancer activity.
25. An isolated nucleic acid molecule selected from:
(a) nucleic acid molecules comprising a nucleotide sequence as shown in SEQ ID NO: 3;
10 (b) nucleic acid molecules comprising a nucleotide sequence capable of hybridizing, under stringent hybridization conditions, to a nucleotide sequence complementary the polypeptide coding region of a nucleic acid molecule as defined in (a) and which codes for a variant form of the *FOXC2* transcription factor; and
(c) nucleic acid molecules comprising a nucleic acid sequence which is degenerate as
15 a result of the genetic code to a nucleotide sequence as defined in (a) or (b) and which codes for a variant form of the *FOXC2* transcription factor.
26. An isolated polypeptide encoded by the nucleic acid according to claim 25.
- 20 27. The isolated polypeptide according to claim 26 having an amino acid sequence shown as SEQ ID NO: 4 in the Sequence Listing
28. A vector harboring the nucleic acid molecule according to claim 25.
- 25 29. A replicable expression vector which carries and is capable of mediating the expression of a nucleotide sequence according to claim 25.
30. A cultured host cell harboring a vector according to claim 28 or 29.
- 30 31. A process for production of a variant form of the *FOXC2* transcription factor polypeptide, comprising culturing a host cell according to claim 30 under conditions whereby said polypeptide is produced, and recovering said polypeptide.

- 24 -

32. A method for identifying an agent capable of regulating expression of the nucleic acid molecule according to claim 25, said method comprising the steps
 - (i) contacting a candidate agent with the said nucleic acid molecule; and
 - (ii) determining whether said candidate agent modulates expression of the said nucleic acid molecule.
- 5
33. An antisense oligonucleotide having a sequence capable of specifically hybridizing to RNA transcribed by the nucleic acid molecule according to claim 25, so as to prevent translation of the said RNA.

10

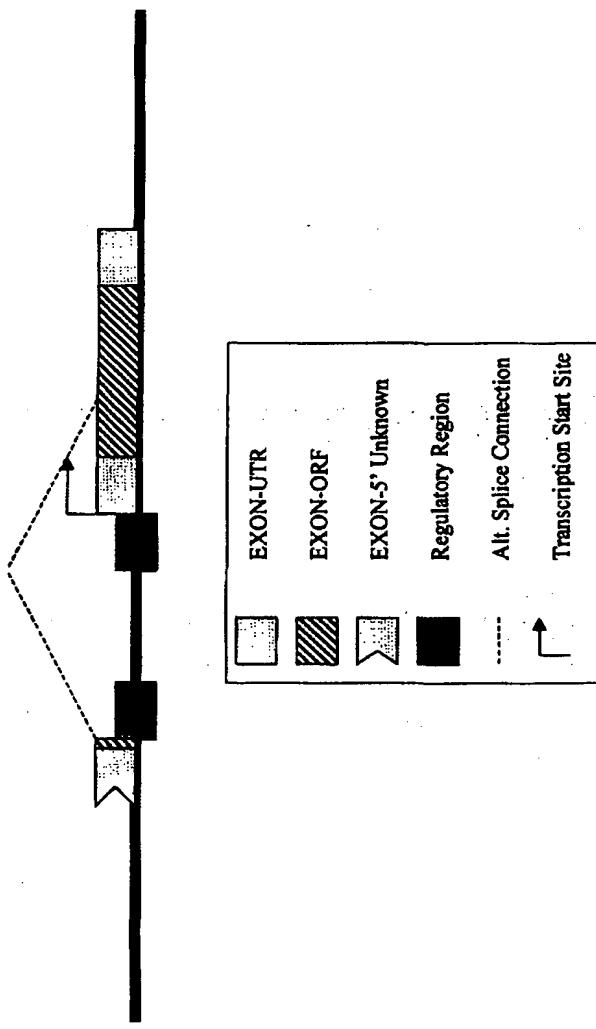
10
11
12
13
14
15
16
17
18
19
20

PRV00-09-26

ABSTRACT

The present invention relates an isolated promoter region of the mammalian transcription factor *FOXC2*. The invention also relates to screening methods for agents modulating the expression of *FOXC2* and thereby being potentially useful for the treatment of medical conditions related to obesity. The invention further relates to a previously unknown variant of the human *FOXC2* gene, derived via the use of an alternative promoter, which produces an additional exon that generates a distinct open reading frame via splicing. The alternative gene encodes a variant of the *FOXC2* transcription factor, which is lacking a part of the DNA-binding domain and consequently has a potential regulatory function.

00000000000000000000000000000000

Fig. 1

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.26.560315>; this version posted September 26, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

PRJEB00409-26

Fig. 2

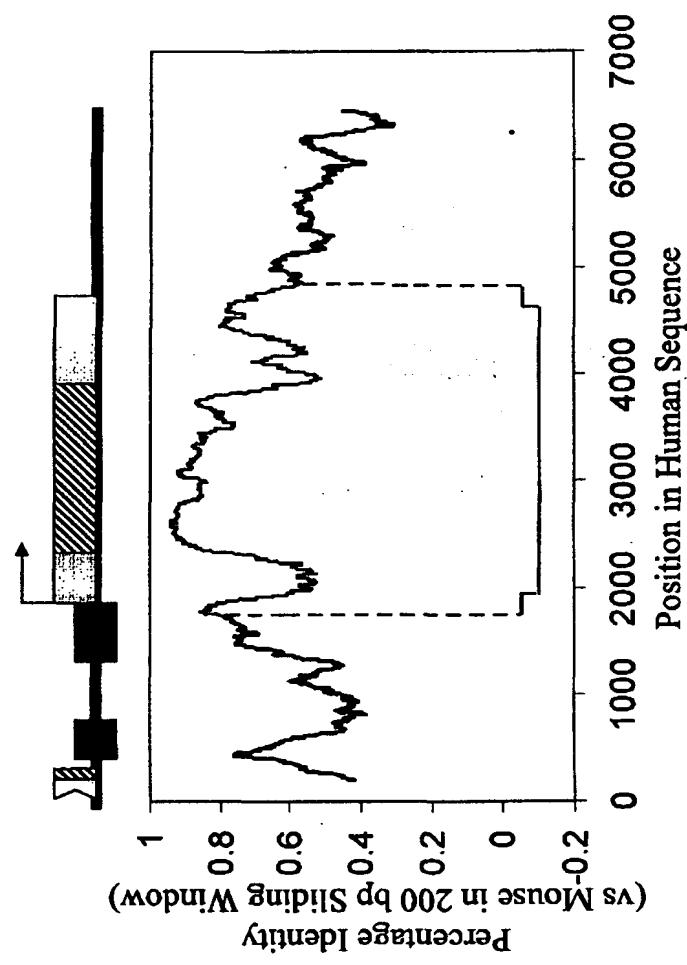


Fig. 3

		Splice	E-box-like	
	200			
HUMAN		AAGGGTGCAGGAAAC <u>ACTGAAATA</u> CACCTGAAGCCCTCTGAAGCCC <u>TCTTCCC</u> AAAC		
MOUSE		CGGAGAAGAAAG <u>ACTGAG</u> --ACA <u>AAATGT</u> TTATCTGTGCCT <u>TCTTCCC</u> AAAC		
HUMAN		CAGACCAGCAACTTCCAAATTCTGCCCGTG--	TTTAGCCTTGTAAAGGGTGTCTCA	
MOUSE		<u>CGGACCAACAA</u> ACTTCCAGGGTCTGCCGAGGCATAGGCCATTCCGTAGGGACATCTCG		
HUMAN		CTCCTTCAGGGAAAGTGGGAAAG--GGGATCTGATT-----	Forkhead-like	
MOUSE		GTGCTTCTGAGGAAGCGGGACCGAGCAGGGATCCGATGACGACTGGAGATGTTGAAGGAAT		
HUMAN		AAATAATCAGTCCACAAATAAACAAACTGTCC--GGGATTCCTAGAGGGAAAGGAATAATC	Ets-like	
MOUSE		<u>AAATA</u> -CCAGTCCACAAATAAACAAACTGTCCCCGGGATTCCTAGAGGGAAAGGAACGCG		
HUMAN		CTTGAAG----GAGATCCAAGTCGCTCCAGGTCTGCCCTGCCAAATAATATC	475	
MOUSE		TTGAAGGTGGGAAACTCCGAGTCGCTGTGGTCAAGGGT-TGGCATAAAATT		

Fig. 4

(1250)	
HUMAN	TGCCATTCCAATC-CAGGGCTTGTCTTGAAATCCATTACACCTGGCCCCATAAATTA
MOUSE	GCCC--TACACGCTCAGTCGGTTGCTC-TGAACCCATTACAACTAGGGCCCGATAAATTA
HUMAN	GGAATCTAATTATCGCTTCACTCATTAATAA-----GAAAATGTCCCAGGAT
MOUSE	AGAAATCTAATTATCGCCTCTCATCCATTAAATAATAAAAAAAATCTCAGGGCT
HUMAN	CATTGCTACTTACAAGGTCTTGGGAGAGATTATTACTCTATTAACTCCATTCTATTAA
MOUSE	CTTTCCTACTTACAAGGTCTGGGGCAAATCTCTGCCAACACTCATCAATTCTGATGTTA
HUMAN	TATTTCAAATTGA-----TTTTTTAAACAGGAAAGTGGCTATC--TTTTTGTGTTG
MOUSE	TATTTCAAACTAAACTCTTTTATTTCACAAAGGAACAGGGTTTTAAATTGCTCTG
HUMAN	GGCATGGGCCATTACAAAATGTGATCATAAAATAAAATTAAATAAGATAACT-
MOUSE	GACACGTGGTCTCGTTAACAAAATGTGATAATAAAATAAAATTAAAGATGTAAC
HUMAN	-TTTAAAGTTCAAGTGAAAGACGGAGTCGC GGAGG-----CCGGGGCGG
MOUSE	ATTTTAAAGTCCTCAAGTTAACCTGAGTCAGTGGCTGGGGGGGGAGATCTGGCTAAAGAGCAT
HUMAN	CGGGGTCTTAGAGCCGACGGATTCCCTGGCTCCCTGCCCGATTGGGCCGGACTCCTCT
MOUSE	CTGGGTCTTAGAGCCGACGGATTCAAGGGCTCCCTGGCTCCCTGGATTGGTGCCTCG
HUMAN	CAGCTGCCGGTGAATTGGCTCAAAGTCCGGAGGGGGCGTGGCCGGCTGAAGAAAGTAAAAA
MOUSE	CAGCTGCCAGATGATTGGTCAAACCTCTGGCTGGGGGGGGGGCTGAAGAAAGTAAAAA
HUMAN	CTCGCTTTCAAGCAAGAACATTGAAACTTTCCCAATCCCTAAAGGGACTTGGCCTC-(1763)
MOUSE	CTCGCTTTGAGCCAGAACATTGAAACTTTCCCAATCCCTAAAGGGACTTGGCTTC

- 1 -

SEQUENCE LISTING

<110> Pharmacia & Upjohn AB

<120> Promoter Sequences

<130> 00298

<140>

<141>

<160> 12

<170> PatentIn Ver. 2.1

<210> 1

<211> 6458

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (2235)..(3740)

<400> 1

cctttggctt tgaattgatc aggagacaaa gataatgcatttgcgtt 60

cttttattgg aaataagtgg cacgccccat tgccttctag tcgcctccca gaagcgaaga 120

ggccgaagcg aagaggcctg gtgggttgc tcaacatcct ttgcgtgaga atcgaatacg 180

cagccgatga acagccagga agggtgcaag gaaacctgaa atacaaatgt tctccctgaa 240

gccctcttcc ctgccaacc agaccagcaa cttccaaat tctgcccgtg tttagccttg 300

ttaaagggt gtctcactcc ttcaggaaa gtggaaaag gggatctgat tattgaggtg 360

tgaaaggaat aaataatcag tccacaaata aacaaactgt ccgggattcc tagagggaaag 420

gagaaatcct tgaaggagat ccaagtcgct ccaggtctgc ctgccaata atatcatccc 480

gaaggatct tgaaccgtt gcaatcaacc gtcacccag tcttccacg gagcgcgc 540

cctaactcac cctacccacc caacaaaaca aaaaaaaggc tgaatatatacg aaaagcaact 600

tggaggctcc cagggggacg ttgccaggag caggaggcag ggacagcgcc ctagggtcgg 660

tgttagcgcc cggcgccggc ctggccacg gaaacgtcc acgcttggtg cccgcggcgc 720

cgccgcgtca ttgcgcgcgc ctgcgacca agccccgcg gaaaacaggc tcgggtttct 780

cctcgcaggg cccaggaact cggctctgcc tggccgggt gggtcgctgc attgtcccg 840

tcttctggga gtgcgggtc agcttggtag agggaatgg tacctggaa aaggagacg 900

agtttcgaag ctgaagttgg taggctgcga gtgtccacgc gggagacgaa agggggaaat 960

agcagagtca cttcacccctt ttccccaaac cccacaaaac tgctcgcagc gacgcggatg 1020

PRU00-09-26

- 2 -

atctaccgaa ttccccgcga attcggagga ttaagttgtc agtcagcacg ttgttaccc 1080
cccccttatg cactccgctg cctggctcct cggcgaaaaa cgaggggaaac tcagttgtc 1140
gggtttaccc ctaaaaacctc gataagggtat ccttgacgc cccgagccct gaaactccct 1200
gttgcattt aattatttga ttaaataagt ataacatcca ggagaggccc tgccattcca 1260
atccagcgcg tttgctttt aatccattac acctggggcc ccataattag gaaatcta 1320
tattcgcttc atcaactcatt aataagaaaa atgtcccagg atcattgcta cttacaaggt 1380
ctttggaga gatattttac tctattaatc cattctattt tatatttcaa attgattttt 1440
tttaacagag gaaagtggct atcttttgt tttggcatg tggggccatt caccaaaatg 1500
tgatcataaa ataaatttta ataagataata actttttaaa aagttttcaa gtgaagacgg 1560
agtcgcccgcg gaggccgggg cggcggggtc ttagagccga cgattccctg cgctccctgc 1620
cccgattggc gccggactcc tctcagctgc cgggtgattt gctcaaagtt ccgggagggg 1680
gcgtggcccg agggaaagtaa aaactcgctt tcagcaagaa gactttgaa actttccca 1740
atccctaaaaa gggacttggc ctcttttctt gggctcagcg gggcagccgc tcggaccccg 1800
gcgcgctgac cctcggggct gccgattcgc tggggcttg gagagcctcc tgcggccctc 1860
ctcgcgcggg ccgaggggtcc accttggtcc ccaggccgcg gcgtctccgc tgggtccgc 1920
gccgcccccc tgcggcgctt gccgcccgcg ggtcctggag ccagcgagga gcggggccgg 1980
cgctgcgtt gcccggggcg cgccctccag gatgccgatc cgcccggtcc gctgaaagcg 2040
cgcccccctg ctccggccga gcgacgacga ccgcgcaccc tcgccccgggaa ggctgccagg 2100
agaccggggc cgccccctccc gctccccctcc tctccccctc tggctctetc gcgctctctc 2160
gctctcaggg ccccccctcgc tccccccggcc gcagtcgtc cgcgagggcg ccggcgagcc 2220
gtctcgaaag cagc atg cag gcg cgc tac tcc gtg tcc gac ccc aac gcc 2270
Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala
1 5 10
ctg gga gtg gtg ccc tac ctg agc gag cag aat tac tac cgg gct gcg 2318
Leu Gly Val Val Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala
15 20 25
ggc agc tac ggc ggc atg gcc agc ccc atg ggc gtc tat tcc ggc cac 2366
Gly Ser Tyr Gly Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His
30 35 40
ccg gag cag tac agc gcg ggg atg ggc cgc tcc tac gcg ccc tac cac 2414
Pro Glu Gln Tyr Ser Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His
45 50 55 60
cac cac cag ccc gcg gcg cct aag gac ctg gtg aag ccg ccc tac agc 2462
His His Gln Pro Ala Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser
65 70 75

- 3 -

tac atc gcg ctc atc acc atg gcc atc cag aac ggc ccc gag aag aag Tyr Ile Ala Leu Ile Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys 80 85 90	2510
atc acc ttg aac ggc atc tac cag ttc atc atg gac cgc ttc ccc ttc Ile Thr Leu Asn Gly Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe 95 100 105	2558
tac cgg gag aac aag cag ggc tgg cag aac agc atc cgc cac aac ctc Tyr Arg Glu Asn Lys Gln Gly Trp Gln Asn Ser Ile Arg His Asn Leu 110 115 120	2606
tcg ctc aac gag tgc ttc gtc aag gtg ccc cgc gac gac aag aag ccc Ser Leu Asn Glu Cys Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro 125 130 135 140	2654
ggc aag ggc agt tac tgg acc ctg gac ccg gac tcc tac aac atg ttc Gly Lys Gly Ser Tyr Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe 145 150 155	2702
gag aac ggc agc ttc ctg cgg cgc cgg cgc ttc aaa aag aag gac Glu Asn Gly Ser Phe Leu Arg Arg Arg Arg Phe Lys Lys Lys Asp 160 165 170	2750
gtg tcc aag gag aag gag gag cgg gcc cac ctc aag gag ccc ccc ccc Val Ser Lys Glu Lys Glu Arg Ala His Leu Lys Glu Pro Pro Pro 175 180 185	2798
gcg gcg tcc aag ggc gcc ccg gcc acc ccc cac cta gcg gac gcc ccc Ala Ala Ser Lys Gly Ala Pro Ala Thr Pro His Leu Ala Asp Ala Pro 190 195 200	2846
aag gag gcc gag aag aag gtg gtg atc aag agc gag ggc ggc tcc ccc Lys Glu Ala Glu Lys Lys Val Val Ile Lys Ser Glu Ala Ala Ser Pro 205 210 215 220	2894
gcg ctg ccg gtc atc acc aag gtg gag acg ctg agc ccc gag agc gcg Ala Leu Pro Val Ile Thr Lys Val Glu Thr Leu Ser Pro Glu Ser Ala 225 230 235	2942
ctg cag ggc agc ccg cgc agc gcg gcc tcc acg ccc gcc ggc tcc ccc Leu Gln Gly Ser Pro Arg Ser Ala Ala Ser Thr Pro Ala Gly Ser Pro 240 245 250	2990
gac ggt tcg ctg ccg gag cac cac gcc ggc ggc ccc aac ggg ctg cct Asp Gly Ser Leu Pro Glu His His Ala Ala Pro Asn Gly Leu Pro 255 260 265	3038
ggc ttc agc gtg gag aac atc atg acc ctg cga acg tcg ccg ccg ggc Gly Phe Ser Val Glu Asn Ile Met Thr Leu Arg Thr Ser Pro Pro Gly 270 275 280	3086
gga gag ctg agc ccg ggg gcc gga cgc gcg ggc ctg gtg ccg ccg Gly Glu Leu Ser Pro Gly Ala Gly Arg Ala Gly Leu Val Val Pro Pro 285 290 295 300	3134

ctg gcg ctg cca tac gcc gcc ggc ccg ccc gcc gcc tac ggc cag ccc Leu Ala Leu Pro Tyr Ala Ala Ala Pro Pro Ala Ala Tyr Gly Gln Pro 305	310	315	3182
tgc gct cag ggc ctg gag gcc ggg gcc gcc ggg ggc tac cag tgc agc Cys Ala Gln Gly Leu Glu Ala Gly Ala Ala Gly Gly Tyr Gln Cys Ser 320	325	330	3230
atg cga gcg atg agc ctg tac acc ggg gcc gag cgg ccg gcg cac atg Met Arg Ala Met Ser Leu Tyr Thr Gly Ala Glu Arg Pro Ala His Met 335	340	345	3278
tgc gtc ccc ccc gcc ctg gac gag gcc ctc tcg gac cac ccg agc ggc Cys Val Pro Pro Ala Leu Asp Glu Ala Leu Ser Asp His Pro Ser Gly 350	355	360	3326
ccc acg tcg ccc ctg agc gct ctc aac ctc gcc gcc ggc cag gag ggc Pro Thr Ser Pro Leu Ser Ala Leu Asn Leu Ala Gly Gln Glu Gly 365	370	375	3374
gcg ctc gcc gcc acg ggc cac cac cac cag cac cac ggc cac cac cac Ala Leu Ala Ala Thr Gly His His His Gln His His Gly His His His 385	390	395	3422
ccg cag gcg ccg ccc ccg ccg gct ccc cag ccc cag ccg acg ccg Pro Gln Ala Pro Pro Pro Pro Ala Pro Gln Pro Gln Pro Thr Pro 400	405	410	3470
cag ccc ggg gcc gcc ggc gcg cag gcg gcc tcc tgg tat ctc aac cac Gln Pro Gly Ala Ala Ala Ala Gln Ala Ala Ser Trp Tyr Leu Asn His 415	420	425	3518
agc ggg gac ctg aac cac ctc ccc ggc cac acg ttc gcg gcc cag cag Ser Gly Asp Leu Asn His Leu Pro Gly His Thr Phe Ala Ala Gln Gln 430	435	440	3566
caa act ttc ccc aac gtg cgg gag atg ttc aac tcc cac cgg ctg ggg Gln Thr Phe Pro Asn Val Arg Glu Met Phe Asn Ser His Arg Leu Gly 445	450	455	3614
att gag aac tcg acc ctc ggg gag tcc cag gtg agt ggc aat gcc agc Ile Glu Asn Ser Thr Leu Gly Glu Ser Gln Val Ser Gly Asn Ala Ser 465	470	475	3662
tgc cag ctg ccc tac aga tcc acg ccg cct ctc tat cgc cac gca gcc Cys Gln Leu Pro Tyr Arg Ser Thr Pro Pro Leu Tyr Arg His Ala Ala 480	485	490	3710
ccc tac tcc tac gac tgc acg aaa tac tga cgtgtccccgg gaccccccct Pro Tyr Ser Tyr Asp Cys Thr Lys Tyr 495	500		3760
ccccggccccg ctccggcttc gcttcccagc cccgacccaa ccagacaatt aaggggctgc 3820			
agagacgcaa aaaagaaaaca aaacatgtcc accaaccttt tctcagacccc gggagcagag 3880			
agcgggcacg ctagccccca gccgtctgtg aagagcgcacg gtaactttaa ttccggccccc 3940			
cgtttctggg atccccaggaa accccctccaa agggacgcacg cccaaacaaaa tgagtattgg 4000			

tcttaaaaatc cccctcccc accaggacgg ctgtgctgtg ctcgacctga gctttcaaaa 4060
gttaagttat ggacccaaat cccatagcga gcccctagtg actttctgta ggggtcccc 4120
taggtgtatg ggggtctcta tagataatat atgtgctgtg tgtaatttta aatttctcca 4180
accgtgctgt acaaatgtgt ggatttgtaa tcaggctatt ttgttgtgt tggtgtgtt 4240
cagagccatt aatataatat ttaaagttga gttcaactgga taagttttc atcttgcaca 4300
accatttcta actgccaaat tgaattcaag aaaccgatgt gggtttgtt tcctgtacaa 4360
ttatgagata taatttttt tccattgtg ggtctttac aaaacaagaa aataatttat 4420
ttttttgttg gtggataaaag aagtcaagta tctgatactt tttatattaca aagtgtgatg 4480
gttttgtata gtaggttcca ccctgagtt tcctaaaaga aaaaaaaaaa aaaagcttaa 4540
aaactctaac ttcatctgtg tttgtcttac gtggctttaa tcgttgtact tacctaaaa 4600
taaacccatg ttgtttttc tgcccaaagt ttggacagtg tgggtgtt gttgcatttt 4660
ttacaaacga ggtgtgttg caaaccacc tgcttgatt attttggta cacaggtggg 4720
tatatgtgt aacacataaa aacgaccaga gaataggagc acacacctgc tgtcttgttt 4780
agtgacagaa aaaggctttt gattaatttt aaaatccac tctaggattt tttctttcg 4840
agaaaccgcc cagttggagg gggctgcctg aaggaccgga ccatgagttt gccgtgatgc 4900
attttcttaa atgcacaaaa acatgcta atgtcaaaaca aacagtgcctt ctccatctca 4960
gtgtccagcc gtccccagtt taggaggtga aggaaggaa gaataaacat ttcccgttt 5020
ctaactgcaa cccagggtaa gtcctgctt ccccgattt tataaaattt gagcctttt 5080
gcctgctta atagtttcc agagaatttgc aactggccca atgaaggctt gaagggacg 5140
gattttctag cgtttgatccatccccct tagcgccag atcagagggg aatttcagac 5200
tttattactt ctcaatgtca tgtctaaatc tacaccctca tcgcagtgaa aaattttaaa 5260
acctcattac cttcaaaaaa taatttatga tatttttaga gttctaaattt caagttttc 5320
aatatgttaa ataatagaga ttattttttgc tttcaatgt taatatctcg tctttacat 5380
ttttatagt aacatagttt ttgtgaaatg tagctgacga aatggcttta ttatctattt 5440
caatggctga agtccaccac tccctgctg gcctctatgt gtgaatttgg ggaccaaagc 5500
ttcatcaattt cccacccag caggtgagct gtaccttgc aatgctgaaatg ttctttgtga 5560
gcttaacgtt tcaagaccag atgatttgc taaaggtgat ttgcttgc gcaatggcgc 5620
tgaacgtaac ccgggtgttt ttgtcggtt gtttcaaca tggcaacttta tctccacgct 5680
atgttgaat agaatttaggg gaagcttaaa gcataataat tgcctccaca tgcgtcaacac 5740

agactcttcc aatctgtggc cccagaggtg gcacacagtgc aagacttggc ggctgtctca 5800
ttcttttca taatgtgcgg gttcccggtt gtccgggtgc tagactttca gcaggccccca 5860
ggccagacgg ctttgggtt agtgaacagg aggaggaagt taaggaggta ggggtgggga 5920
gagaccctct ccaagctgca gaagaaggtg gcccagctc cttgcctgctg tctgccgtga 5980
tggtttcatt ttacttctgc tcgcttcatttgc ctatttgccc caggagaaga ggagagtatt 6040
ccagacggta agcgagctgg cttttcctt tcccttagacg ttttaaagaa atctttctga 6100
aagcttggcc tcacgttaag ctttgaacc gttgggtgtcc tgtagtggc gagggctgag 6160
agacacgcgg agaaataaaag gagagcgcacg gtgtggctga gagccccccag gtctgctgtt 6220
gaaactaagg tgggcttttgc cacctttagg aagcctttt aaagaagtcc tgctgtgtgg 6280
gggccggaaag cccaaagttagt gggccttgtt ggaggttatac gggaggggtc tttaccactc 6340
cttggggaaac gtgggcaacg gggggattgt atctgaagct ttattcaggt cttccggcggc 6400
aqcaqagtgg aqaaccqgc ccttagtqtc taqcggcctg qggattttgg gactcatc 6458

<210> 2
 <211> 501
 <212> PRT
 <213> Homo sapiens

 <400> 2
 Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala Leu Gly Val Val
 1 5 10 15
 Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala Gly Ser Tyr Gly
 20 25 30
 Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His Pro Glu Gln Tyr
 35 40 45
 Ser Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His His His Gln Pro
 50 55 60
 Ala Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu
 65 70 75 80
 Ile Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn
 85 90 95
 Gly Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn
 100 105 110
 Lys Gln Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu
 115 120 125
 Cys Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser
 130 135 140
 Tyr Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser
 145 150 155 160
 Phe Leu Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Ser Lys Glu
 165 170 175
 Lys Glu Glu Arg Ala His Leu Lys Glu Pro Pro Pro Ala Ala Ser Lys
 180 185 190
 Gly Ala Pro Ala Thr Pro His Leu Ala Asp Ala Pro Lys Glu Ala Glu
 195 200 205
 Lys Lys Val Val Ile Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val
 210 215 220

- 7 -

Ile Thr Lys Val Glu Thr Leu Ser Pro Glu Ser Ala Leu Gln Gly Ser
 225 230 235 240
 Pro Arg Ser Ala Ala Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu
 245 250 255
 Pro Glu His His Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val
 260 265 270
 Glu Asn Ile Met Thr Leu Arg Thr Ser Pro Pro Gly Gly Glu Leu Ser
 275 280 285
 Pro Gly Ala Gly Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro
 290 295 300
 Tyr Ala Ala Ala Pro Pro Ala Ala Tyr Gly Gln Pro Cys Ala Gln Gly
 305 310 315 320
 Leu Glu Ala Gly Ala Ala Gly Gly Tyr Gln Cys Ser Met Arg Ala Met
 325 330 335
 Ser Leu Tyr Thr Gly Ala Glu Arg Pro Ala His Met Cys Val Pro Pro
 340 345 350
 Ala Leu Asp Glu Ala Leu Ser Asp His Pro Ser Gly Pro Thr Ser Pro
 355 360 365
 Leu Ser Ala Leu Asn Leu Ala Ala Gly Gln Glu Gly Ala Leu Ala Ala
 370 375 380
 Thr Gly His His His Gln His His Gly His His His Pro Gln Ala Pro
 385 390 395 400
 Pro Pro Pro Pro Ala Pro Gln Pro Gln Pro Thr Pro Gln Pro Gly Ala
 405 410 415
 Ala Ala Ala Gln Ala Ala Ser Trp Tyr Leu Asn His Ser Gly Asp Leu
 420 425 430
 Asn His Leu Pro Gly His Thr Phe Ala Ala Gln Gln Thr Phe Pro
 435 440 445
 Asn Val Arg Glu Met Phe Asn Ser His Arg Leu Gly Ile Glu Asn Ser
 450 455 460
 Thr Leu Gly Glu Ser Gln Val Ser Gly Asn Ala Ser Cys Gln Leu Pro
 465 470 475 480
 Tyr Arg Ser Thr Pro Pro Leu Tyr Arg His Ala Ala Pro Tyr Ser Tyr
 485 490 495
 Asp Cys Thr Lys Tyr
 500

<210> 3
<211> 4158
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (187)...(1437)

<400> 3
cctttggctt tgaattgatc aggagacaaa gataatgcatttccatgttttttcgtttctgttt 60
cttttattgg aaataagtgg cacggcccat tgccttcttag tcgcctcccc gaagcgaaaga 120
ggccgaaagcg aagaggcctg gtgggttgatc tcaacatcct ttgtctgaga atcgaatacg 180
cagccg atg aac agc cag gaa ggg tgc aag gaa acc ttg aac ggc atc 228
Met Asn Ser Gln Glu Gly Cys Lys Glu Thr Leu Asn Gly Ile
1 5 10

- 8 -

tac cag ttc atc atg gac cgc ttc ccc tac cg ^g gag aac aag cag Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn Lys Gln 15 20 25 30	276
ggc tgg cag aac agc atc cgc cac aac ctc tcg ctc aac gag tgc ttc Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu Cys Phe 35 40 45	324
gtc aag gtg ccc cgc gac gac aag aag ccc ggc aag ggc agt tac tgg Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Ser Tyr Trp 50 55 60	372
acc ctg gac ccg gac tcc tac aac atg ttc gag aac ggc agc ttc ctg Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser Phe Leu 65 70 75	420
cgg cgc cgg cgg cgc ttc aaa aag aag gac gtg tcc aag gag aag gag Arg Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Ser Lys Glu Lys Glu 80 85 90	468
gag cgg gcc cac ctc aag gag ccg ccc ccg gcg tcc aag ggc gcc Glu Arg Ala His Leu Lys Glu Pro Pro Ala Ala Ser Lys Gly Ala 95 100 105 110	516
ccg gcc acc ccc cac cta gcg gac gcc ccc aag gag gcc gag aag aag Pro Ala Thr Pro His Leu Ala Asp Ala Pro Lys Glu Ala Glu Lys Lys 115 120 125	564
gtg gtg atc aag agc gag gcg gcg tcc ccg gcg ctg ccg gtc atc acc Val Val Ile Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val Ile Thr 130 135 140	612
aag gtg gag acg ctg agc ccc gag agc gcg ctg cag ggc agc ccg cgc Lys Val Glu Thr Leu Ser Pro Glu Ser Ala Leu Gln Gly Ser Pro Arg 145 150 155	660
agc gcg gcc tcc acg ccc gcc ggc tcc ccc gac ggt tcg ctg ccg gag Ser Ala Ala Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu Pro Glu 160 165 170	708
cac cac gcc gcg ccc aac ggg ctg cct ggc ttc agc gtg gag aac His His Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val Glu Asn 175 180 185 190	756
atc atg acc ctg cga acg tcg ccg ccg ggc gga gag ctg agc ccg ggg Ile Met Thr Leu Arg Thr Ser Pro Pro Gly Gly Glu Leu Ser Pro Gly 195 200 205	804
gcc gga cgc gcg ggc ctg gtg ccg ccg ctg gcg ctg cca tac gcc Ala Gly Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro Tyr Ala 210 215 220	852
gcc gcg ccg ccc gcc gcc tac ggc cag ccg tgc gct cag ggc ctg gag Ala Ala Pro Pro Ala Ala Tyr Gly Gln Pro Cys Ala Gln Gly Leu Glu 225 230 235	900
gcc ggg gcc gcc ggg ggc tac cag tgc agc atg cga gcg atg agc ctg Ala Gly Ala Ala Gly Gly Tyr Gln Cys Ser Met Arg Ala Met Ser Leu 240 245 250	948

- 9 -

tac acc ggg gcc gag cg ^g ccg gc ^g cac atg tgc gtc ccg ccc gcc ctg Tyr Thr Gly Ala Glu Arg Pro Ala His Met Cys Val Pro Pro Ala Leu 255 260 265 270	996
gac gag gcc ctc tcg gac cac ccg agc ggc ccc acg tcg ccc ctg agc Asp Glu Ala Leu Ser Asp His Pro Ser Gly Pro Thr Ser Pro Leu Ser 275 280 285	1044
gct ctc aac ctc gcc ggc cag gag ggc gc ^g ctc gcc gcc acg ggc Ala Leu Asn Leu Ala Ala Gly Gln Glu Gly Ala Leu Ala Ala Thr Gly 290 295 300	1092
cac cac cac cag cac cac gg ^c cac cac ccg cag gc ^g ccg ccc His His His Gln His His His His Pro Gln Ala Pro Pro Pro 305 310 315	1140
ccg ccg gct ccc cag ccc cag ccg acg ccg cag ccc ggg gcc gcc gc ^g Pro Pro Ala Pro Gln Pro Gln Pro Thr Pro Gln Pro Gly Ala Ala Ala 320 325 330	1188
gc ^g cag gc ^g gcc tcc tgg tat ctc aac cac agc ggg gac ctg aac cac Ala Gln Ala Ala Ser Trp Tyr Leu Asn His Ser Gly Asp Leu Asn His 335 340 345 350	1236
ctc ccc ggc cac acg ttc gc ^g gcc cag cag caa act ttc ccc aac gtg Leu Pro Gly His Thr Phe Ala Ala Gln Gln Thr Phe Pro Asn Val 355 360 365	1284
cg ^g gag atg ttc aac tcc cac ccg ctg ggg att gag aac tcg acc ctc Arg Glu Met Phe Asn Ser His Arg Leu Gly Ile Glu Asn Ser Thr Leu 370 375 380	1332
ggg gag tcc cag gtg agt ggc aat gcc agc tgc cag ctg ccc tac aga Gly Glu Ser Gln Val Ser Gly Asn Ala Ser Cys Gln Leu Pro Tyr Arg 385 390 395	1380
tcc acg ccg cct ctc tat cgc cac gca gcc ccc tac tcc tac gac tgc Ser Thr Pro Pro Leu Tyr Arg His Ala Ala Pro Tyr Ser Tyr Asp Cys 400 405 410	1428
acg aaa tac tgacgtgtcc cgggacacctcc cttccccggc ccgtccggc Thr Lys Tyr 415	1477
ttcgcttccc agccccgacc caaccagaca attaaggggc tgcagagacg caaaaaagaa 1537	
acaaaacatg tccacccaacc ttttctcaga cccgggagca gagagcgggc acgctagccc 1597	
ccagccgtct gtgaagagcg caggttaactt taattcgccg ccccgttct gggatcccg 1657	
gaaacccctc caaaggggacg cagcccaaca aaatgagtat tggctttaaa atccccctcc 1717	
cctaccagga cggctgtgct gtgctcgacc ttagcttca aaagttaagt tatggaccca 1777	
aatcccatag cgagcccta gtgactttct ttaggggtcc ccataagggtgt atgggggtct 1837	
ctataagataa tatatgtgct gtgtgtaatt ttaaatttct ccaaccgtgc tgtacaaatg 1897	

- 10 -

tgtggatttg taatcaggct attttgttgt tggtgttgc gtcagagcc attaatataa 1957
tattnaaagt tgagttcaact ggataagttt ttcatcttgc ccaaccattt ctaactgcc 2017
aattgaattc aagaaaccga tgggggtttt gtttcctgta caattatgag atataattct 2077
ttttcccatt gtaggtcttt tacaaaacaa gaaaataatt tatttttttgc ttgggtggata 2137
aagaagtcaa gtatctgata ctttttattt acaaagtgtg atgggtttgt atagtaggtt 2197
ccaccctgag tattcctaaa agaaaaaaaaaaaaaaagct taaaaactct aacttcatct 2257
gtgttgtct tacgtggtct taatcggtt acttaccta aaataaaccc atgttgttt 2317
ttctgcacaa agttggaca gtgtgtttgt gttgttgcatttttacaa cgaggtgtgt 2377
ttgcaaaccc acctgcttg attatTTTtgc ttacacaggt ggttatatgt gtagacacat 2437
aaaaacgacc agagaatagg agcacacacc tgctgtcttgc tttagtgaca gaaaaaggct 2497
tttgattaat tttaaaatcc cactcttagga tttttcttt tcgagaaacc gcccagttgg 2557
agggggctgc ctgaaggacc ggaccatgag tttgcctgta tgcattttct taaatgcaca 2617
aaaacatgct aattgtcaaa acaaacagtgc ccactccatc tcagtgtcca gccgtcccc 2677
gtttaggagg tgaaggaagg gaagaataaa catttccgt ttgcttaactg caacccaggg 2737
tgagtccctgc ttcccccgat tttataaaaa ttgagccctc ttgcctgct ttaatagttt 2797
tccagagaat ttgaactggg ccaatgaagg tctgaagggg acggattttc tagcgttgc 2857
tatccatccc ccttagccgc cagatcagag gggaaattca gacttttata cttctcaatg 2917
tcatgtctaa atctacaccc tcatcgcatg gaaaaattttt aaaacctcat tacccttcaa 2977
aaataattta tgatattttt agagttctaa attcaagttt ttcaatatgt taaataatag 3037
agattatttt ttgtttcaa tggtaatatac tcgtcttttgcattttat agtaacatag 3097
tttttgtgaa atgttagctga cgaaatggct ttattatcta ttcaatggc tgaagtccac 3157
cactccctg ctggccctcta tggtaatttgcggacaa agttcatca attccaccc 3217
cagcaggtga gctgtacctt gctaattgtcg aagttcttgc ttgagcttac gtttcaagac 3277
cagatgatt tgctaaagggt gatTTTgcattt gatgcagtgg cgctgaacgt aacccgggtg 3337
ttttgtcgt gttgtttca acatggcact ttatctccac gctatgtga aatagaattt 3397
ggggaaagctt aaagcataat aattgtcccc acatgtgcaaa cacagactct ttcaatctgt 3457
ggccccagag gtggcacaca gttaaagactt ggcggctgtc tcattttttt tcataatgtg 3517
cggttcccg ggtgtccggg tgcttagactt tcagcaggcc ccaggccaga cgggcttgg 3577
ttgagtgaac aggaggagga agttaaggag gttaggggtgg ggagagaccc tctccaagct 3637
gcagaagaag gtggcccaag ctccctgcct gcgtctgccc tgatggtttc attttacttc 3697

- 11 -

tgctcgcttc atgctatttgc ccccaggaga agaggagagt attccagacg gtaagcggc 3757
 tggcttttc cttcccttag acgtttaaa gaaatcttc tgaaagcttg ccctcatcgt 3817
 aagctttgaa accgttggtg tcctgttagt ggcgaggcgt gagagacacg cggagaaata 3877
 aaggagagcg acgggtgtggc tgagagcccc caggctctgct gttgaaacta agctggcctt 3937
 ttgcacctt aggaagcctt tttaaagaag tcctgtgtg tgggggcccgg aagcccaagt 3997
 gagtgggcct tgtggagggtt atcgggaggg gtcttacca ctcttgggg aacgtggca 4057
 acgggggat tgtatctgaa gcttattca ggtctcggc ggcagcagag tggagaacca 4117
 ggcccttagt gtgtacggc ctggggattt tggactcat c 4158

<210> 4
 <211> 417
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Asn Ser Gln Glu Gly Cys Lys Glu Thr Leu Asn Gly Ile Tyr Gln
 1 5 10 15

Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn Lys Gln Gly Trp
 20 25 30

Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu Cys Phe Val Lys
 35 40 45

Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser Tyr Trp Thr Leu
 50 55 60

Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser Phe Leu Arg Arg
 65 70 75 80

Arg Arg Arg Phe Lys Lys Asp Val Ser Lys Glu Lys Glu Arg
 85 90 95

Ala His Leu Lys Glu Pro Pro Ala Ala Ser Lys Gly Ala Pro Ala
 100 105 110

Thr Pro His Leu Ala Asp Ala Pro Lys Glu Ala Glu Lys Lys Val Val
 115 120 125

Ile Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val Ile Thr Lys Val
 130 135 140

Glu Thr Leu Ser Pro Glu Ser Ala Leu Gln Gly Ser Pro Arg Ser Ala
 145 150 155 160

Ala Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu Pro Glu His His
 165 170 175

Ala Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val Glu Asn Ile Met
 180 185 190

- 12 -

Thr Leu Arg Thr Ser Pro Pro Gly Gly Glu Leu Ser Pro Gly Ala Gly
 195 200 205
 Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro Tyr Ala Ala Ala
 210 215 220
 Pro Pro Ala Ala Tyr Gly Gln Pro Cys Ala Gln Gly Leu Glu Ala Gly
 225 230 235 240
 Ala Ala Gly Gly Tyr Gln Cys Ser Met Arg Ala Met Ser Leu Tyr Thr
 245 250 255
 Gly Ala Glu Arg Pro Ala His Met Cys Val Pro Pro Ala Leu Asp Glu
 260 265 270
 Ala Leu Ser Asp His Pro Ser Gly Pro Thr Ser Pro Leu Ser Ala Leu
 275 280 285
 Asn Leu Ala Ala Gly Gln Glu Gly Ala Leu Ala Ala Thr Gly His His
 290 295 300
 His Gln His His Gly His His His Pro Gln Ala Pro Pro Pro Pro Pro
 305 310 315 320
 Ala Pro Gln Pro Gln Pro Thr Pro Gln Pro Gly Ala Ala Ala Ala Gln
 325 330 335
 Ala Ala Ser Trp Tyr Leu Asn His Ser Gly Asp Leu Asn His Leu Pro
 340 345 350
 Gly His Thr Phe Ala Ala Gln Gln Gln Thr Phe Pro Asn Val Arg Glu
 355 360 365
 Met Phe Asn Ser His Arg Leu Gly Ile Glu Asn Ser Thr Leu Gly Glu
 370 375 380
 Ser Gln Val Ser Gly Asn Ala Ser Cys Gln Leu Pro Tyr Arg Ser Thr
 385 390 395 400
 Pro Pro Leu Tyr Arg His Ala Ala Pro Tyr Ser Tyr Asp Cys Thr Lys
 405 410 415

Tyr

<210> 5
 <211> 6021
 <212> DNA
 <213> Mus musculus

 <220>
 <221> exon
 <222> (1649)..(4348)

 <300>
 <308> GenBank/Y08222

<309> 1997-05-14

<300>

<301> Miura, N
<303> Genomics
<304> 41
<306> 489-492
<307> 1997

<400> 5

ctcgagtcaa aggttagcaca cataaaaacctt attttgcgtc ttccggtaacgt caagcaatgc 60
cactaaagggtt tcctcacccg ccaaagctga aacagtgggt tctaattctt caaaggcttt 120
tgccgaaaat ctaaagggggg tggggggctt tggtgggtgc gtgggggggg ggtcgagaa 180
gaagaaaagac tgagacaaat gtttatctg tcgccttcc ccttaccctaa ccggaccaac 240
aacttccaga aggttctgcg aggcatagaa ccattccgtt ggacatctc ggtgcttctg 300
aggaagcggg ccgagcaggg atccgatgac gactggagat gtgaaaggaa taaataccag 360
tccacaaaata aacaactgt ccccggtt cctagaggaa agagacacgc ttgaaggctg 420
ggaaactccg agtcgtgtc cgtaagggtt ggcataaaaat taaaaaaaaaaa aaaagtccctt 480
cagttaccag gcccctctaag gagccccctgg tcctcaagtc accttatcaa aactcagtaa 540
aacaacacgc ctgaataaca gtcattttac aggatccaa agatgctgac cgccggatgg 600
gaccacgcgccc gggcccccggc aacagctggg gaagccgggc cgaggctaca cagtccgcg 660
ctcccttgcg tttccagtgtc cgaagccggc gatggagtgc aggttggag ctccccacgc 720
cgaacggggg caccagctcc cgggggtctt ctgccttgc ctaacctcca gacagcgctt 780
tcataggtgg ggagaaggga gaggccggg tggatggcag gggaaagctg ccctcgctt 840
tgcgggagag gagaccaggaa aagcaacagt tgggttcaacg cgcttccctg aaccccacga 900
aattgtttgg aggactcaga tggatcacct aagtacgacg gaagacgaag gaccaatgg 960
tccttaggtt ttacccccc agttggcat tcccactaag cttccctcc cagcccgacc 1020
ccgtcgtaa gggggagggg accgaattct ccaaccggc ctcccttgc ggctcttcct 1080
caacctggaa gcgtctgtt aattatccat cactgcattt aacaggccct acacgctcag 1140
tccgtttgtt ctgaacccat tacaactagg ccccgataat taagaatct aattatcg 1200
ctcttcatcc attaataata ataaaaaaaaaa aatctccagg ctcttcctt cttacaagg 1260
cttggggca aatctgtcc caacttcattt aattcgatgt tatatttcaa actaaacttc 1320
tttttatttt ccaaaggaaac agggtttttta attttgcgtc tggacacgtg gtctcgtaa 1380
acaaaatgtt ataataaaat aaatattttt aagatgtaac tcatattttaa aagtccctaa 1440
gttaacttga gctggggggg gggggagatct ggtaagagc atctgggtct tagacccgac 1500
ggattcaggc getccctcggtt ttgatgggtt ccatccctt cgcagctgac agatgattgg 1560
tgcaaaacttc ctggggggg cgcggcctga agaaaagtaaa aactcgctt gagccagaag 1620
acttttggaa cttttcccaa tccctaaaag ggactttgtt ttttttccg ggctcgccg 1680
cgccgcctt cgggacccctt agetcggtac gctcgccgtt gcaatctcc tggcggggcc 1740
cgagagccgc tgcgtctttt ttcggactt ggaagggtt gtgcgtctt acggtcgcgc 1800
gtggcgctgt tgccgccagc tcagggtctt caccggccaa gcccggccat cgcggccagc 1860
ggggccgcctt gccgtcacc cttcaggatg cccatccggc cggtcggctg aacccgagcg 1920
ccggcgctt ccgcgcgtgg accgcgaggc tggcccgatg cggggctgac tgcacgtctc 1980
cgcccccttc tgctctctt ctccggccctt cgcgcgcgc gggccgcgtt cggtcgcgc 2040
aggcggcgac cgggcgtctt ggacgcgcac tgcaggcgat ttaactcggtt tggacccca 2100
acgccttggg agtggatccc tatttggatg agaaaaacta ctaccggccg gccggcagct 2160
acggcgccat ggccagcccc atgggcgtctt actccggcca cccggagcag tacggcgccg 2220
gcatggggccg ctccctacgcg ccctaccacc accagccgcg ggcggcccaag gacctggta 2280
agccgcctt cagctatata ggcgtcatca ccatggcgat ccagaacgcg ccagagaaga 2340
agatcactt gaacggcattt taccaggttt tcattggaccg tttcccttc taccgcgaga 2400
acaaggcaggc ctggcagaac agcatccgc acaaccgtt actcaatggat tgcttcgtga 2460
aagtgcggcgcc cgacgacaag aagccggccg agggcageta ctggacgcgc gacccggact 2520
cctacaacat gttcgagaat ggcagcttcc tgcggccggcg gccggcgatcc aagaagaagg 2580
atgtgccccaa ggacaaggag gagccggccccc acctcaaggg gccccttc accacggcca 2640
agggcgctcc gacaggacc cccgttagctt acggggccaa ggaggccgag aagaaagtctg 2700
tggtaagag cgaggccgcg tcccccgcgc tgccggctcat caccacggatg gagacgttgc 2760
ggcccgaggg agcgtcgcag gccagtcgc gcagcgcattt ctccacggccc gcagggttccc 2820
cagacggctt gctgcggag caccacggccg cggccctaa cgggtcgccc ggcttcagcg 2880
tggagaccat catgacgctg cgcacgtcgc ctccggccgg cgatctgagc ccagcggccg 2940

cgcgccgg cctgggtgg ccacccgtgg cactgccata cgccgcagcg ccacccggc 3000
 cttacacgca gcccgtcg cagggcctgg aggctgcggg ctccgcggg taccagtgc 3060
 gtatgcgggc tatagtcgt tacacccggg cccgagcggcc cgcgcacgtg tgcgttccgc 3120
 cccgcgtgga cgagctctg tcggaccacc cgagcggccc cgctccccg ctgcggcggc 3180
 tcaacctcgc agcgggtcag gagggcgcgt tggggcctc ggttaccac caccagcatc 3240
 acggccacct ccacccgcag ggcgcacccg cccgcgcga gccccctccc ggcgcgcagc 3300
 cccgcaccca ggccacctcc tggtatctga accacccggg ggacctgagc cacctccccg 3360
 gccacacgtt tgcaacccaa cagcaaactt tcccaacgt ccgggagatg ttcaactcgc 3420
 accggctagg actggacaac tgcgtccctcg gggagtccca ggtgagcaat gcgagctgtc 3480
 agctgccta tgcgactacg cgcgtccctc accgcacgc agccccctac ttttacact 3540
 gcaccaaata ctgaggctgt ccagtcgcgt ccagccccag gaccgcacccg gttcgcctc 3600
 ctccatggga accttctcg acggagccgc agaaagcgac gaaagcgcc cctctctcag 3660
 aaccaggaga agagagctc gtgcacactcg caggttaactt atccgcagct cagtttgaga 3720
 tctcagcgag tccctctaag ggggatgcag cccagaaaa cggaaatacag atttttttt 3780
 taattccttc ccctacccag atgcgtgcgc tgctccctt gggcttcata gattagctt 3840
 tggaccaaac ccatagggac ccctaattgac ttctgtggag attctccacg ggcgcagag 3900
 gtctctccgg ataaggtgcc ttctgtaaac gagtgcggat ttgttaaccag gctattttgt 3960
 tcttgcggc agccttaat ataataattt aagtgtgtc cactggataa gtttgcgtct 4020
 tgcccaactg ttactgccaa attgaattca agaaacgtgt gttttgcgtt tctcccaacg 4080
 tcaccatgat aaaataggc cctcccaaaa ctgttaggtt ttacaaaac aagaaaataa 4140
 tttatTTTT tttttttttt tgtagttgtt ggataacgaa attaagtatc ggataactttt aatTTTggaa 4200
 gtgcgttgtt ttgtacagta gatgcctatc ggggtattcc aaaaacacac caaaagactt 4260
 taaaatttca atctcacctg tttttgtt atgtgatctc agtggatgtt ttacctaaa 4320
 ataaaacccgt gtttttttcc tggccaaatg tcggacagag tctttgtgtt cttaattttt 4380
 aaaaggggaaa ttgttagtaag ccagttgtga ttgtatTTT tgatgcagg tggcctggta 4440
 acgtggatgc atatacaggt tacaggacga tggagctctc gatttagtaat agaaggggct 4500
 cttgattttt tgaactatcc cgtccctgata tttttttttt ttctgcgtca ggttaatctga 4560
 gaaactgttc tccatccaca cacggacagg gtcgcctgag gcaacacgtcc tgctggctg 4620
 ttaacgaaat gcttgcggg atgcagaaaaa ctgttccaa ttgtcaaaaac aaaaatgtgt 4680
 caccctgtt cgggttccag ctgtccctcg tttaggggaa gaaaccgaga aaggacaaac 4740
 ggcctgcagc ttgcttaacct cagcttagca ggagcctggg tgagtgcgtc gctccctca 4800
 tttcctttaga tgccgacttg ttgcctctgt tggcggttta agagtgcacg caagaagcaa 4860
 agagggttgg taggtctctg gtatTTTact gcccgttgg ggttcgatc agaagtgtat 4920
 ttcagtctga ttatTTTTTTT aattttggct ttaaatattt tactccggcg tgggtggaaaa 4980
 agaagccact gtgcgcctcc agcatgatattttttttt tttttttttt tttttttttt 5040
 atgctaagta acaggagatt atttttttttt tgattttttt atttcatttc tttttttttt 5100
 aaaaaggaaa tagatcgaaa caaactctct aaaaatgtacc tgctggctg ggggtgggtc 5160
 cttaccaatc tgctgcctga aagatacagc ttcagcacag gctgcgtgt tggactttag 5220
 gcatatcatc gattcccacg ccagttggta acctggactg tctaatggaa agttctttct 5280
 gcacagaaca tgtagggccag gaggaggcag ggacccggg ggggggtggta ctttgcagg 5340
 catctgttta gcttagtggt gcccacgggt taacacgtat atgtgttac tgggtggaaac 5400
 tccaagttt atatctgtgc tttttttttt tagaattttgg ggaggttccct gatgatacta 5460
 ccctacccgt gtatgtaaa cagtcttca acctgcgtg ccagaatgtg acccacactt 5520
 cagttatctc cataaaagtgg ggggactaaag aactggacag ggttgcgtg gagggggggca 5580
 ggcctgggtt atcttgggtt ctgagcagag cagagatctt agaagggggt cgggagatct 5640
 ctgggttccctc ccaacactgg tttttttt catggctctc ttccaaacctc ttggccctagg 5700
 agaagcgagc ttgttccaaag ccagttggctt ccgttccctt ccagatgtt tagggggctc 5760
 cctgaaagct tgccctccctc ttaagattca gaactccgtt cccaggaaaa gataggaggc 5820
 tttgtggatg ggatTTTTT tttaaagagg accgttctcg ttctcaagta ggttagctaga 5880
 gagaagcccc ctggagcagg ccctacttgc gactgtcagg gaaccaggat tttttttttt 5940
 gctttccca ggcctcccaag agcagcgggt tgaaaaatg cgttccctggg aaaagtgggt 6000
 ctgggggtttt gcttccctcg g 6021

<210> 6
 <211> 2712
 <212> DNA
 <213> Mus musculus

- 15 -

<220>
<221> CDS
<222> (422)..(1906)

<300>
<308> GenBank/Y08222
<309> 1997-05-14

<300>
<301> Miura, N
<303> Genomics
<304> 41
<306> 489-492
<307> 1997

<400> 6
agggactttg cttcttttc cgggctcggc cgcgacgcct ctccggaccc tagctcgctg 60
acgctgcggg ctgcagttct cctggcgaaa cccgagagcc gctgtctctt tttctagcac 120
tcggaagggc tgggtgcgt ccacggtcgc gcgtggcgcc tggccgcca gctcagggt 180
gccacccgccc aagccgagag tgccggccca gcggggccgc ctgcccgtca cccttcagga 240
tgccgatccg cccggctggc tgaacccgag cgccggcgctt ttccgcgcgt ggaccgcgag 300
gctgccccga gtcggggctg cctgcacatcgc tccgtccctt cctgctctcc tgctccgggc 360
ctcgctcgcc gcggggccgca gtcgggtgcgc gcaggccgc accggggcgcc tgggacgcag 420
c atg cag gcg cgt tac tcg gta tcg gac ccc aac gcc ctg gga gtg gta 469
Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala Leu Gly Val Val
1 5 10 15

ccc tat ttg agt gag caa aac tac tac cgg ggc gcc ggc agc tac ggc 517
Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala Gly Ser Tyr Gly
20 25 30

ggc atg gcc agc ccc atg ggc gtc tac tcc ggc cac ccg gag cag tac 565
Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His Pro Glu Gln Tyr
35 40 45

ggc gcc ggc atg ggc cgc tcc tac gcg ccc tac cac cac cag ccc gcg 613
Gly Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His His Gln Pro Ala
50 55 60

gcg ccc aag gac ctg gtg aag ccg ccc tac agc tat ata gcg ctc atc 661
Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile
65 70 75 80

acc atg gcg atc cag aac gcg cca gag aag atc act ctg aac ggc 709
Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn Gly
85 90 95

atc tac cag ttc atc atg gac cgt ttc ccc ttc tac cgc gag aac aag 757
Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn Lys
100 105 110

cag ggc tgg cag aac agc atc cgc cac aac ctg tca ctc aat gag tgc Gln Gly Trp Gln Asn Ser Ile Arg His Asn Lys Ser Leu Asn Glu Cys 115 120 125	805
ttc gtg aaa gtg ccg cgc gac gac aag aag ccg ggc aag ggc agc tac Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Ser Tyr 130 135 140	853
tgg acg ctc gac ccg gac tcc tac aac atg ttc gag aat ggc agc ttc Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser Phe 145 150 155 160	901
ctg cgg cgg cgg cgg cgc ttc aag aag aag gat gtg ccc aag gac aag Leu Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Pro Lys Asp Lys 165 170 175	949
gag gag cgg gcc cac ctc aag gag ccg ccc tcg acc acg gcc aag ggc Glu Glu Arg Ala His Leu Lys Glu Pro Pro Ser Thr Thr Ala Lys Gly 180 185 190	997
gct ccg aca ggg acc ccg gta gct gac ggg ccc aag gag gcc gag aag Ala Pro Thr Gly Thr Pro Val Ala Asp Gly Pro Lys Glu Ala Glu Lys 195 200 205	1045
aaa gtc gtg gtt aag agc gag gcg gcg tcc ccc gcg ctg ccc gtc atc Lys Val Val Val Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val Ile 210 215 220	1093
acc aag gtg gag acg ctg agc ccc gag gga gcg ctg cag gcc agt ccg Thr Lys Val Glu Thr Leu Ser Pro Glu Gly Ala Leu Gln Ala Ser Pro 225 230 235 240	1141
cgc agc gca tcc tcc acg ccc gca ggt tcc cca gac ggc tcg ctg ccc Arg Ser Ala Ser Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu Pro 245 250 255	1189
gag cac cac gcc gcg cct aac ggg ctg ccc ggc ttc agc gtg gag Glu His His Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val Glu 260 265 270	1237
acc atc atg acg ctg cgc acg tcg cct ccg ggc gat ctg agc cca Thr Ile Met Thr Leu Arg Thr Ser Pro Pro Gly Gly Asp Leu Ser Pro 275 280 285	1285
gcg gcc gcg cgc gcc ggc ctg gtg cca ccg ctg gca ctg cca tac Ala Ala Ala Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro Tyr 290 295 300	1333
gcc gca gcg cca ccc gcc gct tac acg cag ccg tgc gcg cag ggc ctg Ala Ala Ala Pro Pro Ala Ala Tyr Thr Gln Pro Cys Ala Gln Gly Leu 305 310 315 320	1381
gag gct gcg ggc tcc gcg ggc tac cag tgc agt atg cgg gct atg agt Glu Ala Ala Gly Ser Ala Gly Tyr Gln Cys Ser Met Arg Ala Met Ser 325 330 335	1429
ctg tac acc ggg gcc gag ccg ccc gcg cac gtg tgc gtt ccg ccc gcg Leu Tyr Thr Gly Ala Glu Arg Pro Ala His Val Cys Val Pro Pro Ala 340 345 350	1477

ctg gac gag gct ctg tcg gac cac ccg agc ggc ccc ggc tcc ccg ctc	1525
Leu Asp Glu Ala Leu Ser Asp His Pro Ser Gly Pro Gly Ser Pro Leu	
355 360 365	
ggc gcc ctc aac ctc gca gcg ggt cag gag ggc gcg ttg ggg gcc tcg	1573
Gly Ala Leu Asn Leu Ala Ala Gly Gln Glu Gly Ala Leu Gly Ala Ser	
370 375 380	
ggt cac cac cac cag cat cac ggc cac ctc eac ccg cag ggc cca ccg	1621
Gly His His His Gln His Gly His Leu His Pro Gln Ala Pro Pro	
385 390 395 400	
ccc gcc ccg cag ccc cct ccc gcg ccg cag ccc gcc acc cag gcc acc	1669
Pro Ala Pro Gln Pro Pro Pro Ala Pro Gln Pro Ala Thr Gln Ala Thr	
405 410 415	
tcc tgg tat ctg aac cac ggc ggg gac ctg agc cac ctc ccc ggc cac	1717
Ser Trp Tyr Leu Asn His Gly Gly Asp Leu Ser His Leu Pro Gly His	
420 425 430	
acg ttt gca acc caa cag caa act ttc ccc aac gtc cgg gag atg ttc	1765
Thr Phe Ala Thr Gln Gln Thr Phe Pro Asn Val Arg Glu Met Phe	
435 440 445	
aac tcg cac ccg cta gga ctg gac aac tcg tcc ctc ggg gag tcc cag	1813.
Asn Ser His Arg Leu Gly Leu Asp Asn Ser Ser Leu Gly Glu Ser Gln	
450 455 460	
gtg agc aat gcg agc tgt cag ctg ccc tat cga gct acg ccg tcc ctc	1861
Val Ser Asn Ala Ser Cys Gln Leu Pro Tyr Arg Ala Thr Pro Ser Leu	
465 470 475 480	
tac cgc cac gca gcc ccc tac tct tac gac tgc acc aaa tac tga	1906
Tyr Arg His Ala Ala Pro Tyr Ser Tyr Asp Cys Thr Lys Tyr	
485 490 495	
ggctgtccag tccgctccag cccccaggacc gcacccggctt cgccctccctcc atggaaacct	1966
tcttcgacgg agccgcagaa agcgacggaa agcgccctc tctcagaacc aggagcagag	2026
agctccgtgc aactcgcagg taacttatcc gcagctcagt ttgagatctc agcgagtc	2086
tctaaggggg atgcagccca gcaaaacgaa atacagattt tttttttat tccttccct	2146
acccagatgc tgccctgtct cccttggggc ttcatagatt agcttatgga ccaaaccat	2206
agggaccctt aatgacttct gtggagattc tccacggcg caagaggctt ctccggataa	2266
ggtgccctct gtaaaacgagt gcggtttgt aaccaggcta ttttgttctt gcccagagcc	2326
ttaatataa tatttaaatgt tttgtccact ggataagggtt tttgtttgcc caactgttac	2386
tgccaaattt aattcaagaa acgtgtgtgg gtctttctc cccacgtcac catgataaaa	2446
taggtccctc cccaaactgt aggtctttta caaaaacaaga aaataattt tttttttgtt	2506
gttggat aacgaaatta agtacggat acttttaatt taggaagtgc atggcttgc	2566

- 18 -

acagtagatg ccatctgggg tattccaaaa acacacccaa agactttaaa atttcaatct 2626
 cacctgtgtt tgtcttatgt gatctcagtg ttgtatttac cttaaaataa acccggtttg 2686
 ttttctgcc caaaaaaaaaa aaaaaa 2712

<210> 7
 <211> 494
 <212> PRT
 <213> Mus musculus

<400> 7
 Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala Leu Gly Val Val
 1 5 10 15
 Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala Gly Ser Tyr Gly
 20 25 30
 Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His Pro Glu Gln Tyr
 35 40 45
 Gly Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His His Gln Pro Ala
 50 55 60
 Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile
 65 70 75 80
 Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn Gly
 85 90 95
 Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn Lys
 100 105 110
 Gln Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu Cys
 115 120 125
 Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser Tyr
 130 135 140
 Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser Phe
 145 150 155 160
 Leu Arg Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Pro Lys Asp Lys
 165 170 175
 Glu Glu Arg Ala His Leu Lys Glu Pro Pro Ser Thr Thr Ala Lys Gly
 180 185 190
 Ala Pro Thr Gly Thr Pro Val Ala Asp Gly Pro Lys Glu Ala Glu Lys
 195 200 205
 Lys Val Val Val Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val Ile
 210 215 220
 Thr Lys Val Glu Thr Leu Ser Pro Glu Gly Ala Leu Gln Ala Ser Pro
 225 230 235 240
 Arg Ser Ala Ser Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu Pro
 245 250 255
 Glu His His Ala Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val Glu
 260 265 270
 Thr Ile Met Thr Leu Arg Thr Ser Pro Pro Gly Gly Asp Leu Ser Pro
 275 280 285
 Ala Ala Ala Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro Tyr
 290 295 300
 Ala Ala Ala Pro Pro Ala Ala Tyr Thr Gln Pro Cys Ala Gln Gly Leu
 305 310 315 320
 Glu Ala Ala Gly Ser Ala Gly Tyr Gln Cys Ser Met Arg Ala Met Ser
 325 330 335
 Leu Tyr Thr Gly Ala Glu Arg Pro Ala His Val Cys Val Pro Pro Ala
 340 345 350
 Leu Asp Glu Ala Leu Ser Asp His Pro Ser Gly Pro Gly Ser Pro Leu
 355 360 365

- 19 -

Gly	Ala	Leu	Asn	Leu	Ala	Ala	Gly	Gln	Glu	Gly	Ala	Leu	Gly	Ala	Ser
370							375					380			
Gly	His	His	His	Gln	His	His	Gly	His	Leu	His	Pro	Gln	Ala	Pro	Pro
385							390					395			400
Pro	Ala	Pro	Gln	Pro	Pro	Pro	Ala	Pro	Gln	Pro	Ala	Thr	Gln	Ala	Thr
							405					410			415
Ser	Trp	Tyr	Leu	Asn	His	Gly	Gly	Asp	Leu	Ser	His	Leu	Pro	Gly	His
							420					425			430
Thr	Phe	Ala	Thr	Gln	Gln	Gln	Thr	Phe	Pro	Asn	Vai	Arg	Glu	Met	Phe
							435					440			445
Asn	Ser	His	Arg	Leu	Gly	Leu	Asp	Asn	Ser	Ser	Leu	Gly	Glu	Ser	Gln
							450					455			460
Val	Ser	Asn	Ala	Ser	Cys	Gln	Leu	Pro	Tyr	Arg	Ala	Thr	Pro	Ser	Leu
465							470					475			480
Tyr	Arg	His	Ala	Ala	Pro	Tyr	Ser	Tyr	Asp	Cys	Thr	Lys	Tyr		
							485					490			

<210> 8
<211> 3289
<212> DNA
<213> Homo sapiens

<300>
<301> Miura, N
<303> Genomics
<304> 41
<306> 489-492
<307> 1997

<300>
<308> GenBank/Y08223
<309> 1997-05-14

<400> 8
gaattcggag gattaagttg tcagtcagca cgttgtacc ttccctcta tgcactccgc 60
tgcctggctc ctcggcgaaa agcgaggaa actcagttt tagggttac ctctaaaacc 120
tcgatagtt atccttgacg acccccggcc tggaaactcc ctgttgatga ttaattattt 180
gattaaataa gtataacatc caggagaggg cctgcattc caatccagcg cgtttgcctt 240
tgaatccatt acacccggc cccataattt agggaaatcta attattcgct tcataactca 300
ttaataagaa aaatgtccca ggttcatgtt tacttacaag gtcggggga gagatatttt 360
actctattaa tccatcttat tttatattt aaattgtttt ttttaacag agggaaatgg 420.
ctatctttt gttttggca tttggggccca ttccaaaaaa ttgtatccata aaataaattt 480
taataagata taactttta aaaagttttc aagttaagac ggatcgccg cggaggccgg 540
ggccggcgaaa tcttagagcc gacggattcc tgcgttcctc gccccgattt ggcggact 600
cctctcact gccgggtgat tggctcaaaat ttccgggggg gggcgtggcc cgaggaaat 660
aaaaactcgc ttccagcaag aagactttt aaactttcc caatccctaa aaggacttg 720
gcctctttt ctgggttcag cggggcagcc gtcggaccc cggcgccgtt accctcgaaa 780
ctggccatc gctgggggct tggagagcc cctgcgtcc tccctcgccg gcccgggggt 840
ccacccgtt cccaggccg cggcgatcc gctgggtccg cggccggcccg cctggcccg 900
ctggccggcc cgggttccgg agccagcgag gggcgccgc ggcgtgcgc ttggccgggg 960
cgccgcctcc aggatgcga tccggccggc cgcgttccaa cgcgcgcgc tgcgtggccc 1020
gaggcgacgac gaccggcgcac ctcggccccc gaggctgcca ggagaccggg gccgcgcctc 1080
ccgttccctt cctctccccc tctggatctc tgcgttcctc tgcgttcctc ggcggccctc 1140
gtccccccgg ccggcgttccg tggcgaggg cggccggccg cgcgttcggaa agcagcatgc 1200
aggcgccgtt ctcgggttcc gaccccaacg ccctgggggt ggtgcctac ctggcgagc 1260
agaattacta ccgggttccg ggcgttccg gggcatggc cagccccatg ggcgttattt 1320
ccggccaccc ggagcgttcc acgcgcgggg tggccgttc ctacgcgc taccaccacc 1380

- 20 -

accagccgc ggcgcctaag gacctggta agccgccta cagctacatc gcgctcatca 1440
 ccatggccat ccagaacgcg cccgagaaga agatcacett gaacggcatc taccagtta 1500
 tcatggaccc cttccccttc taccggaga acaaggagg ctggcaqaac agcatccgccc 1560
 acaaccttc gctcaacgag tgcttcgtca aggtccccg cgacgacaag aagccggca 1620
 agggcagttt ctggaccctg gacccggact cctacaacat gtgcgagaac ggcagttcc 1680
 tgccggccg gcggcgttc aaaaagaagg acgtgtccaa ggagaaggag gagcgggccc 1740
 acctaaga gcccggccg gccgcgtcca agggcgcccc gcccaccccc cacctagcg 1800
 acgcccccaa ggaggccgag aagaagggtg tgatcaagag cgaggccgccc tccccggcgc 1860
 tgccggcat caccaagggtg gagacgctga gccccggagag cgcgctgcag ggcagccgc 1920
 gcaacgcggc ctccacgcggc gccggctccc cgcacgggtc gtcgcggag caccacgc 1980
 cggcgcccaa cgggcgtcct ggcttcagcg tggagaacat catgaccctg cgaacgtcgc 2040
 cgccggggcg agagctgagc cggggggccg gacgcgcggg cttgggtgtg cccgcgtgg 2100
 cgctgccata cgccgcgcg cggccggccg cctacggcca gccgtgcgtc caggccctgg 2160
 aggccggggc cgccgggggc taccagtgcg gcatgcgagc gatgagcctg tacaccggg 2220
 cccagccggc ggcgcacatg tgcgtccgc cgcgcgttca cggccctggc cgaggccctc tcggaccacc 2280
 cgagccggccc cacgtcgccc ctgagcgctc tcaacccgcg cggccggccag gagggccgc 2340
 tcgcccacac gggccaccac caccagcacc acggccacca ccaccgcag gcgcgcgc 2400
 ccccgccggc tccccagccc cagccgacgc cgcagccgg gggccgcgcg ggcgcaggcgg 2460
 cctcttgta tctcaaccac agcggggacc tgaaccacat ccccgccac acgttcgcgg 2520
 cccagcaga aactttcccc aacgtgcggg agatgtcaa tccccaccgg ctggggattt 2580
 agaactcgac ctcggggag tcccgaggta gtggcaatgc cagctgcag ctgcctaca 2640
 gatccacgccc gcctcttat cgcacgcgg cccctactc ctacgactgc acgaaatact 2700
 gacgtgtccc gggacccccc ctcccccggc cgctccggc tcgccttcca gccccgaccc 2760
 aaccagacaa ttaaggggct gcagagacgc aaaaaagaaa caaaacatgt ccaccaacct 2820
 ttttctcagac cccggagcag agagccggca cgctagcccc cagccgtctg tgaagagcgc 2880
 aggttaacttt aattccgcgc cccgtttctg ggatccagg aaaccctcc aaaggagacgc 2940
 agcccaacaa aatgagtatt ggtttaaaaa tccccctccc ctaccaggac ggctgtgtc 3000
 tgctcgacct gagtttcaa aagttaagtt atggacccaa atcccatagc gagcccttag 3060
 tgactttctg taggggtccc catagggtta tgggggtctc tatagataat atatgtgtc 3120
 tgttaattt taaatttctc caaccgtgtc gtacaaatgt gtggatttgc aatcaggctc 3180
 ttttgttgtt gttgtgttgc ttcaagagcca ttaatataat attaaagtt gagttactg 3240
 gataagttt tcatcttgcgca accatttc taactgcacca attgaattc 3289

<210> 9
 <211> 1506
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1506)

<300>
 <308> GenBank/NM_005251
 <309> 1999-12-23

<400> 9
 atg cag gcg cgc tac tcc gtg tcc gac ccc aac gcc ctg gga gtg gtg 48
 Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala Leu Gly Val Val
 1 5 10 15

ccc tac ctg agc gag cag aat tac tac cgg gct gcg ggc agc tac ggc 96
 Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala Gly Ser Tyr Gly
 20 25 30

ggc atg gcc agc ccc atg ggc gtc tat tcc ggc cac ccg gag cag tac 144
 Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His Pro Glu Gln Tyr
 35 40 45

agc gcg ggg atg ggc cgc tcc tac gcg ccc tac cac cac cac cag ccc Ser Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His His His Gln Pro	192
50 55 60	
gcg gcg cct aag gac ctg gtg aag ccg ccc tac agc tac atc gcg ctc Ala Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu	240
65 70 75 80	
atc acc atg gcc atc cag aac gcg ccc gag aag aag atc acc ttg aac Ile Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn	288
85 90 95	
ggc atc tac cag ttc atc atg gac cgc ttc ccc ttc tac cgg gag aac Gly Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn	336
100 105 110	
aag cag ggc tgg cag aac agc atc cgc cac aac ctc tcg ctc aac gag Lys Gln Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu	384
115 120 125	
tgc ttc gtc aag gtg ccc cgc gac gac aag aag ccc ggc aag ggc agt Cys Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser	432
130 135 140	
tac tgg acc ctg gac ccc gac tcc tac aac atg ttc gag aac ggc agc Tyr Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser	480
145 150 155 160	
ttc ctg cgg cgc cgg cgg cgc ttc aaa aag aag gac gtg tcc aag gag Phe Leu Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Ser Lys Glu	528
165 170 175	
aag gag gag cgg gcc cac ctc aag gag ccg ccc ccg gcg tcc aag Lys Glu Glu Arg Ala His Leu Lys Glu Pro Pro Pro Ala Ala Ser Lys	576
180 185 190	
ggc gcc ccc gcc acc ccc cac cta gcg gac gcc ccc aag gag gcc gag Gly Ala Pro Ala Thr Pro His Leu Ala Asp Ala Pro Lys Glu Ala Glu	624
195 200 205	
aag aag gtg gtg atc aag agc gag gcg gcg tcc ccg ccg ctg ccg gtc Lys Lys Val Val Ile Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val	672
210 215 220	
atc acc aag gtg gag acg ctg agc ccc gag agc gcg ctg cag ggc agc Ile Thr Lys Val Glu Thr Leu Ser Pro Glu Ser Ala Leu Gln Gly Ser	720
225 230 235 240	
ccg cgc agc gcg gcc tcc acg ccc gcc ggc tcc ccc gac ggt tcg ctg Pro Arg Ser Ala Ala Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu	768
245 250 255	
ccg gag cac cac gcc gcg ccc aac ggg ctg cct ggc ttc agc gtg Pro Glu His His Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val	816
260 265 270	

PRV00-09-26

- 22 -

- 23 -

<210> 10
<211> 501
<212> PRT
<213> Homo sapiens

<400> 10
Met Gln Ala Arg Tyr Ser Val Ser Asp Pro Asn Ala Leu Gly Val Val
1 5 10 15
Pro Tyr Leu Ser Glu Gln Asn Tyr Tyr Arg Ala Ala Gly Ser Tyr Gly
20 25 30
Gly Met Ala Ser Pro Met Gly Val Tyr Ser Gly His Pro Glu Gln Tyr
35 40 45
Ser Ala Gly Met Gly Arg Ser Tyr Ala Pro Tyr His His His Gln Pro
50 55 60
Ala Ala Pro Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu
65 70 75 80
Ile Thr Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn
85 90 95
Gly Ile Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn
100 105 110
Lys Gln Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu
115 120 125
Cys Phe Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser
130 135 140
Tyr Trp Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser
145 150 155 160
Phe Leu Arg Arg Arg Arg Phe Lys Lys Lys Asp Val Ser Lys Glu
165 170 175
Lys Glu Glu Arg Ala His Leu Lys Glu Pro Pro Pro Ala Ala Ser Lys
180 185 190
Gly Ala Pro Ala Thr Pro His Leu Ala Asp Ala Pro Lys Glu Ala Glu
195 200 205
Lys Lys Val Val Ile Lys Ser Glu Ala Ala Ser Pro Ala Leu Pro Val
210 215 220
Ile Thr Lys Val Glu Thr Leu Ser Pro Glu Ser Ala Leu Gln Gly Ser
225 230 235 240
Pro Arg Ser Ala Ala Ser Thr Pro Ala Gly Ser Pro Asp Gly Ser Leu
245 250 255
Pro Glu His His Ala Ala Ala Pro Asn Gly Leu Pro Gly Phe Ser Val
260 265 270
Glu Asn Ile Met Thr Leu Arg Thr Ser Pro Pro Gly Gly Glu Leu Ser
275 280 285
Pro Gly Ala Gly Arg Ala Gly Leu Val Val Pro Pro Leu Ala Leu Pro
290 295 300
Tyr Ala Ala Ala Pro Pro Ala Ala Tyr Gly Gln Pro Cys Ala Gln Gly
305 310 315 320
Leu Glu Ala Gly Ala Ala Gly Gly Tyr Gln Cys Ser Met Arg Ala Met
325 330 335
Ser Leu Tyr Thr Gly Ala Glu Arg Pro Ala His Met Cys Val Pro Pro
340 345 350
Ala Leu Asp Glu Ala Leu Ser Asp His Pro Ser Gly Pro Thr Ser Pro
355 360 365
Leu Ser Ala Leu Asn Leu Ala Ala Gly Gln Glu Gly Ala Leu Ala Ala
370 375 380
Thr Gly His His His Gln His His Gly His His His Pro Gln Ala Pro
385 390 395 400

Pro Pro Pro Ala Pro Gln Pro Gln Pro Thr Pro Gln Pro Gly Ala
 405 410 415
 Ala Ala Ala Gln Ala Ala Ser Trp Tyr Leu Asn His Ser Gly Asp Leu
 420 425 430
 Asn His Leu Pro Gly His Thr Phe Ala Ala Gln Gln Gln Thr Phe Pro
 435 440 445
 Asn Val Arg Glu Met Phe Asn Ser His Arg Leu Gly Ile Glu Asn Ser
 450 455 460
 Thr Leu Gly Glu Ser Gln Val Ser Gly Asn Ala Ser Cys Gln Leu Pro
 465 470 475 480
 Tyr Arg Ser Thr Pro Pro Leu Tyr Arg His Ala Ala Pro Tyr Ser Tyr
 485 490 495
 Asp Cys Thr Lys Tyr
 500

<210> 11
 <211> 327
 <212> DNA
 <213> Homo sapiens

<300>
 <308> GenBank/AW271272
 <309> 2000-01-03

<300>
 <301> Strausberg, Robert
 <303> Trends Genet.
 <304> 16
 <305> 3
 <306> 103-106
 <307> 2000

<400> 11
 ttttttttac attttcgtct tctgttcttg tgattggaaa taagtggcac gccccattgc 60
 cttctagtcg cctcccccga gcgaaaggcc cgaagcgaag aggcttgggt ggttgtctca 120
 acatcccttt gctgagaatc gaatacgcag ccgatgaaca gccaggaagg gtgcaggaa 180
 accttgaacg gcatctacca gttcatcatg gaccgcttcc cttcttaccg ggagaacaag 240
 cagggctggc agaacagcat cccgcacaac ctctcgctca acgagtgtt cgtcaagggtg 300
 ccccgcgacg acaagaagcc cggcaag 327

<210> 12
 <211> 147
 <212> DNA
 <213> Homo sapiens

<300>
 <308> GenBank/AW793237
 <309> 2000-05-16

<300>
 <301> Dias Neto, E
 <303> Proc. Natl. Acad. Sci. U.S.A.
 <304> 97
 <306> 3491-3496
 <307> 2000

-25-

<400> 12

ccgtctgaga atcgaatacg cagccgatga acagccagga agggtgcaag gaaaccttga 60
acggcatcta ccagttcata atggaccgct tcccccttcta ccgggagaac aagcaggct 120
ggcagaacag catccgcccac aacctct 147